

A-77N: Ed Hart.

Access DB# 119047

## SEARCH REQUEST FORM

Scientific and Technical Information Center

CRFE

Requester's Full Name Celine Qian Examiner # 78710 Date: 4/8/04  
An Unit 1636 Phone Number 306-2-0777 Serial Number 101090458  
Mail Box and Bldg Room Location 2A87 Results Format Preferred (circle) PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc., if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention Novel ABCAS Transporter and uses thereof  
Inventors (please provide full names): Chen, Hongyu et al.

Earliest Priority Filing Date 3/1/2001

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search SEQ ID NO: 4, also do an oligo search of SEQ ID NO: 4  
Also search a nucleic acid molecule encoding SEQ ID NO: 5.

### STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher _____	NA Sequence (#) <u>2</u> STN _____	
Searcher Phone # _____	AA Sequence (#) <u>1</u> - <u>request to NIA</u>	
Searcher Location _____	Structure (#) _____	Quoted Other _____
Date Searcher Provided <u>4/8/04</u>	Bibliographic _____	Or Other _____
Date Submitted <u>4/15/04</u>	Citation _____	Index News _____
Searcher Prep & Review Time _____	Fulltext _____	Sequence Systems <u>76H</u>
Client Prep Time _____	Patent Family _____	WWW Internet _____
Other Time _____	Other _____	Other _____



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 119047**

**TO: Celine Qian**  
**Location: rem/2a89/2c70**  
**Art Unit: 1636**  
**Thursday, April 15, 2004**

**Case Serial Number: 10/090458**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**REM-1A55**  
**Phone: 571-272-2512**

**[edward.hart@uspto.gov](mailto:edward.hart@uspto.gov)**

### **Search Notes**

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart





# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor  
571-272-2507 Remsen E01 D86

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library Remsen Bldg.



Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches. Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions .nrpm and .nrpn  
Searches run against the Amino Acid Pending database produce two sets of results, with the extensions .rapm and .rapn

*Because they contain data that is confidential, the results of Pending database searches should not be left in the case .*

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2004, 00:54:04 ; Search time 252 Seconds  
(without alignments)  
12056.969 Million cell updates/sec

Title: US-10-090-458-4  
Perfect score: 5475  
Sequence: 1 gcgtccgcgcctcgacag.....cataagcaatgtgaaagtt 5475

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	760.2	13.9	974	4	US-09-833-381-1085
2	319.2	5.8	485	4	US-09-621-976-409
3	319.2	5.8	1350	4	US-09-833-381-1082
4	159.2	2.9	5894	3	US-08-665-259-24
5	159.2	2.9	5894	3	US-08-762-500-24
6	159.2	2.9	5525	3	US-08-762-500-74
7	137.6	2.5	7860	4	US-09-525-193A-2
8	103.4	1.9	1064	4	US-08-859-207A-88
9	97.4	1.8	728	4	US-09-833-381-1087
10	92.2	1.7	5625	4	US-08-961-527-1
11	83.8	1.5	450	4	US-09-621-976-2582
12	83.4	1.5	463	4	US-09-621-976-2581
13	83.4	1.5	478	4	US-09-621-976-2583
14	80.8	1.5	465	4	US-08-833-381-1080
15	80.6	1.5	795	4	US-08-956-171E-264
16	77.8	1.4	8536	4	US-08-956-171E-278
17	77.2	1.4	696	4	US-09-134-001C-987
18	75.8	1.4	1026	4	US-09-134-001C-1205
19	75	1.4	774	4	US-09-107-532A-412
20	74.6	1.4	1664976	4	US-08-916-421B-1
21	72.2	1.3	768	4	US-09-107-532A-1681
22	71	1.3	1664976	4	US-08-916-421B-1
23	69.8	1.3	990	4	US-09-107-532A-1190
24	69.8	1.3	28626	4	US-09-596-002-15
25	69.2	1.3	969	4	US-09-107-532A-1008
26	68	1.2	7963	4	US-08-956-171E-168
27	67.4	1.2	1599	4	US-09-543-681A-474

c	28	67	1.2	527	4	US-09-833-381-1083	Sequence 1083, Ap
c	29	66.8	1.2	19124	2	US-08-487-826B-13	Sequence 13, Appl
c	30	66.2	1.2	714	4	US-09-134-000C-1821	Sequence 1821, Ap
c	31	65.8	1.2	4848	4	US-08-961-527-185	Sequence 185, App
c	32	65.8	1.2	5717	4	US-08-961-527-102	Sequence 102, App
c	33	65	1.2	1770	4	US-09-328-352-3466	Sequence 3466, Ap
c	34	64.8	1.2	1785	4	US-09-543-681A-1356	Sequence 1356, Ap
c	35	64.6	1.2	3144	4	US-08-961-527-224	Sequence 224, Appl
c	36	64.6	1.2	6273	4	US-08-961-527-21	Sequence 21, Appl
c	37	64	1.2	16397	4	US-08-956-171E-205	Sequence 205, App
c	38	63.2	1.2	4093	4	US-08-956-171E-108	Sequence 108, App
c	39	63	1.2	353	4	US-09-833-381-384	Sequence 384, App
c	40	63	1.2	14672	4	US-08-961-527-111	Sequence 111, App
c	41	62.6	1.1	711	4	US-09-134-000C-1433	Sequence 1433, Ap
c	42	62.6	1.1	7218	1	US-08-232-463-14	Sequence 14, Appl
c	43	62.4	1.1	1554	4	US-09-134-000C-1110	Sequence 1110, Ap
c	44	62.4	1.1	580073	4	US-08-545-528D-1	Sequence 1, Appl
c	45	62	1.1	2163	4	US-09-134-000C-2226	Sequence 2226, Ap

## ALIGNMENTS

### RESULT 1

US-09-833-381-1085

; Sequence 1085, Application US/09833381

; Patent No. 6672186

; GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.

; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

; FILE REFERENCE: 5800-119

; CURRENT APPLICATION NUMBER: US/09/833,381

; PRIOR FILING DATE: 2001-04-11

; PRIOR APPLICATION NUMBER: 09/516,448

; NUMBER OF SEQ ID NOS: 2050

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1085

; LENGTH: 974

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-833-381-1085

Query Match 13.9%; Score 760.2; DB 4; Length 974;  
Best Local Similarity 99.0%; Pred. No. 8.6e-168;  
Matches 765; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy	83	CGGCTCGGCGCCCTGAGGTTTATTTCAGAAAACATGTCCTGCAATTAGGGAGTAGGAGT	142
Db	202	CTGCCCCCGCCACAGGTTTATTTCAGAAAACATGTCCTGCAATTAGGGAGTAGGAGT	261
Qy	143	TTGGAGACAGACGACAGACACTTCTACTGAGAAATTAATTAATGAGAACCAAAA	202
Db	262	TTGGAGACAGACGACAGACACTTCTACTGAGAAATTAATTAATGAGAACCAAAA	321
Qy	203	GAGTAGTGTTTCAGGAAATCTTTTCCACTATTTTATTTTGGTTAATTAATTAG	262
Db	322	GAGTAGTGTTTCAGGAAATCTTTTCCACTATTTTATTTTGGTTAATTAATTAG	381
Qy	263	CATGATGCATCCAAATAGAAATATAGAAAGTCCCTAATATAGAACTCAATCCCTATGA	322
Db	382	CATGATGCATCCAAATAGAAATATAGAAAGTCCCTAATATAGAACTCAATCCCTATGA	441
Qy	323	CAAGTTTACTCTTTCTTAATCTAATCTTCCAGTACTCCAGTGACTAATATTACAGCAG	382
Db	442	CAAGTTTACTCTTTCTTAATCTAATCTTCCAGTACTCCAGTGACTAATATTACAGCAG	501
Qy	383	CATCATGGAAGAGTGCTTACTGATCATCTGATGTCATTAATTAATGAGAAATATAC	442
Db	502	CATCATGGAAGAGTGCTTACTGATCATCTGATGTCATTAATTAATGAGAAATATAC	561
Qy	443	AAATGAAAAGAAATGTTTAACTCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGTGT	502



```

QY 4625 TATCATGCTGTCTGGCGAGTTAAGATGATTCGGAACAGTCAACATCTAAGAGTAATT 4684
DB 645 CATCATGCTGTCTGGAGGCTTAGATGATTTGGCTCCATCCACACCTGAAACAACT 704
QY 4685 TGGAAAGGCTACTTTTGGAAATTAATTAAGAGCTGGATAGAAAACCTTAGAAGTAGA 4744
DB 705 TGGCAAGGATTACATCTTAGAGCTAAAGTGAAG-----GAAAGCTCTCAAGTGAC 755
QY 4745 CGCGCTTCAAGAGAAATTCAGTATATTTCCCAANTGCCAAGCGCTCAGGAAGTTTTC 4804
DB 756 TTTGTCACACTGAGATCTGAGAGCTTTTCCACAGGCTGCGAGGAGGAAGGTATTC 815
QY 4805 TTCTATTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCAAAATCTTTTT 4864
DB 816 CTCCTTTGTAACCTATAAGCTGCCGTGGCAGAGCTTTTACCCTCTATCACAGACCTTCA 875
QY 4865 TAAGCTGAAGAGCTAAACATGCTTTTGGCAATTAAGATATAGCTTTTCTCAGCAAC 4924
DB 876 CAATTAAGACAGTGAAGCATAACTTTTAACTCGAAGATACAGCCCTTTCTCAGTGAC 935
QY 4925 ATTGGAACAGCTTTTGTAGAACTCACTTAAAGAAACAAGAGGAGGAAGATTAAGTTGGG 4984
DB 936 ACTGAGAGAGTATCTTAGAGCTTTCTAAGAACAGGAGTAGGAATTTTGTATGAGA 995
QY 4985 AACT 4988
DB 996 AATT 999

```

RESULT 4

```

US-08-665-259-24
; Sequence 24, Application US/08665259
; Patent No. 6028173
; GENERAL INFORMATION:
; APPLICANT: Landes, Gregory M.
; APPLICANT: Burn, Timothy C.
; APPLICANT: Connors, Timothy D.
; APPLICANT: Dackowski, William R.
; APPLICANT: Van Raay, Terence J.
; APPLICANT: Klinger, Katherine W.
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,
; METHODS OF MAKING AND USING SAME
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENZYME CORPORATION
; STREET: One Mountain Road
; CITY: Framingham
; STATE: Massachusetts
; COUNTRY: United States of America
; ZIP: 01701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,259
; FILING DATE: 17-JUN-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dugan, Deborah A.
; REGISTRATION NUMBER: 37,315
; REFERENCE/DOCKET NUMBER: IGS-9.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 872-8400
; TELEFAX: (508) 872-5415
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..5053
; US-08-665-259-24

Query Match 2.9%; Score 159.2; DB 3; Length 5894;
Best Local Similarity 48.0%; Pred. No. 2.4e-27;
Matches 554; Conservative 0; Mismatches 588; Indels 12; Gaps 3;

QY 1317 ATGCTCACACTTAATAGATATATTCATGTCCTTGGCTGTCTATCTTGATCAAGTCAAT 1376
DB 1304 ATGCTGCTCTGCACTCTGTGCTCTATGGCTGTGACCTGTGTACATGAGGCCGCTTTC 1363
QY 1377 CCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCAATATGGTCA 1436
DB 1364 CCAGGGCAGTTCGGGTGCTCAGCCCTGGTACTTCTTCATCATGCCCCCTCTATTTGGT 1423
QY 1437 AAGAGCAAAAGAAATTAAGAGGATTAAGAGGCAATGTTAATGGAATAATTTAGTTTT 1496
DB 1424 GGGAAAGCCAAAGGGCGGTTCAGGGAAGGAGGAGACAGTGAACCCCGAGAAAGCCTC 1483
QY 1497 AGTGAATTTATTGAGCCAGTTCCTTTCAGAAATTTGTAGGAAGAAAGAGCCATAAGAT 1556
DB 1484 AAGAACGAGTACTTTGAAGCCGAGCCAGGAGCTGG---TGGCGGGATCAAGTCAAG 1540
QY 1557 GGTATTTCAGAGACATACAGAAAGAGGTGAAATGTGAGGCTTTGAGAAATTTGTCA 1616
DB 1541 CACCTGTCCAAGGTGTTTCAAGGGTGGAAATAAGGACAGGGCGGCGCTCAGAGACCTGA 1600
QY 1617 TTTGACATATATAGGGTCAGATTAATCTGCTTACTTGGCCACAGTGGAAACAGGAAGCT 1676
DB 1601 CTCACCTGTACGAGGAGCAGATCACCGTCTCTGCTGGGCGCAACAGCGTCCCGGAGACC 1660
QY 1677 ACATTGATGAATATTTCTTGTGGAATCTGCCACCTCTCTGATGGGTTCATCTATATAT 1736
DB 1661 ACCACCTCTCCATGCTACAGGTCTCTTTCCCGCCACAGTGGAGCGGCAATACATCAG 1720
QY 1737 GGACACAGAGTCTCAGAAATAGATGAATTTGTTGAGCAAGAAATGATTTGGCATTTGT 1796
DB 1721 GGGTATGAATTTCCCGAG-----GACATGGTTCAGATCCCGAAGAGGCTGGGCTGTGC 1774
QY 1797 CCACAGTTAGATATACACTTTGATGTTTTCAGAGTAGAAGAAATTTATCAATTTGGCT 1856
DB 1775 CCGCAGCAGCATCTCTGTTTGAACACTTTCAGTTCGACAGACCTTTATTTCTACGCC 1834
QY 1857 TCAATCAAGAGGATACAGCCCAACATATAATACAGAAAGTGCAGAGGTTTATAGAT 1916
DB 1835 CAGCTGAAGGGCTCTGTCAGTCAAGAGTGCCTGGAAGTCAAGCAGATGCTGCACATC 1894
QY 1917 TTAGACATGCACACTATCAAGATTAACCAAGCTAAATAATTAAGTGGTGGTCAAAAAGA 1976
DB 1895 ATCGGCTGGAGACAAGTGGAACTCAGGAGCCGCTTCTGAGCGGGGCAATGAGGCGC 1954
QY 1977 AAGCTGTCAATTAGGAATTTGCTTCTTGGGAACCCAAAGATACCTGCTCTAGATGAACA 2036
DB 1955 AAGCTCTCCATCGCATCGCCCTCATCGCAGGCTCCAAGTGTCTGATACTGACGAGGCC 2014
QY 2037 ACAGCTGGAAATGGACCCCTGTTCTCGACATTTGATGGAATCTTTTAAATACAGAAA 2096
DB 2015 ACTCGGCGCATGGACGCCATCTCCAGGAGGGCCATCTGGGATCTTCTTCAGCGGAGAA 2074
QY 2097 GCCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTTCAGAT 2156
DB 2075 AGTGACCGCACCATCTGTGCTGACCCACCCACTTTCATGGACGAGGCTGACCTGCTGG 2134
QY 2157 AGGAAGCTGTGATATACAGGAATGCTGAATGTGTTGTTCTTCAATGTTCTCTCAAA 2216
DB 2135 CGCATCGCCATCATCGTCCCAAGGGGAGTGTGAGTCTGCGGGTCTCTGCTGTTCTCAAG 2194
QY 2217 AGTAAATGGGATCGGCTACCGCTGAGCATGTATACATAGACAAATATTTGTGCCACAGA 2276
DB 2195 CAGAAATACGGTGGCGGCTATCATATGACGCTGGTGGAGGAGCCGACCTGCAACCCGAA 2254

```



Db 2315 GCCGAGCTGCTTTCATCCTCCAGAGAGAGCAGCAGGTTTGAAGGTTCTTTGCT 2374  
 QY 2397 GCCCTAGACAGTCANTCAAAATTTGGTGTCTATTTC---TTAAGTGTTTCATGACCACT 2453  
 Db 2375 AAACCTGGAGAGAGAGCAGAGAGAGCTGGCATTCGCCATTTGGGGCATCCATCACCACC 2434  
 QY 2454 TTGGAAGACGTATT 2467  
 Db 2435 ATGGAGGAAGTCTT 2448

RESULT 6

US-08-762-500-74  
 ; Sequence 74, Application US/08762500  
 ; Patent No. 6030806  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Landes, Gregory M.  
 ; APPLICANT: Burn, Timothy C.  
 ; APPLICANT: Connors, Timothy D.  
 ; APPLICANT: Dackowski, William R.  
 ; APPLICANT: Van Raay, Terence J.  
 ; APPLICANT: Klinger, Katherine W.  
 ; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES,  
 ; TITLE OF INVENTION: COMPOSITIONS, METHODS OF MAKING AND USING SAME  
 ; NUMBER OF SEQUENCES: 83  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: GENZYME CORPORATION  
 ; STREET: One Mountain Road  
 ; CITY: Framingham  
 ; STATE: Massachusetts  
 ; COUNTRY: United States of America  
 ; ZIP: 01701  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/762.500  
 ; FILING DATE: 09-DEC-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/665,259  
 ; FILING DATE: 17-JUN-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/10469  
 ; FILING DATE: 17-JUN-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Dugan, Deborah A.  
 ; REGISTRATION NUMBER: 37,315  
 ; REFERENCE/DOCKET NUMBER: IGS-9.3  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (508) 872-8400  
 ; TELEFAX: (508) 872-5415  
 ; INFORMATION FOR SEQ ID NO: 74:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 6525 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 573..5684  
 ; US-08-762-500-74

Query Match 2.9%; Score 159.2; DB 3; Length 6525;  
 Best Local Similarity 48.0%; Pred. No. 2.5e-27;  
 Matches 554; Conservative 0; Mismatches 588; Indels 12; Gaps 3;  
 QY 1317 ATGCTCACACTTAATAGTATATTCTTCTTGGCTGTCTATCTTGATCAAGTCATT 1376

1935 ATGCTGCTGCTGGAGCTCTGTGCTCTATGGCTGGTGACCTGGTACATGGAGGCCGCTTTC 1994  
 QY 1377 CAGGGGGAATTTGGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCATATATGGTCA 1436  
 Db 1995 CCAGGCGAGTTGGGCTGCTCAGCCCTGGTACTTTCTTCATGCCCCCTCTATTGGTGT 2054  
 QY 1437 AGAGCAAAAGAAATTTATGAGGAGTTATCAGAGGGCAATGTTAAATGGAATATTAGTTTT 1496  
 Db 2055 GGGAGCCCAAGGCGGTTGCAGGAGGAGGAGAGACAGTACCCCCGAGAAAGCACTC 2114  
 QY 1497 AGTGAATATTAGAGCCAGTTTCTTCAGAAATTTGTAGAAAAGAGCCCATAGAAATAGT 1556  
 Db 2115 AGAAACGAGTACTTTTGAAGCCGAGCCAGAGGACCTGG---TGGCGGGATCAGATCAAG 2171  
 QY 1557 GGTATTTCAGAGACATACAGAAAGAGGCTGAAATGTGAGGCTTTGAGAAATTTGTCA 1616  
 Db 2172 CACCTGTCCAAGGTGTTCAAGGTGGGAATAGGACAGGCGCGCGCTCAGAGACCTGAAC 2231  
 QY 1617 TTGACATATATGAGGCTCAGATTACTGCTTACTTGGCCACAGAGTGGAAACAGAAAGT 1676  
 Db 2232 CTCACCTGTACGAGGACAGATCACCGTCTCTGCTGGGCCACACAGCGTCCCGGAGACC 2291  
 QY 1677 ACATTGATGAATATTCTTTTGTGACTCTGCCACCTTCTGTATGGGTTTTCATCTATAT 1736  
 Db 2292 ACCACCTCTCCATGCTCAGGCTCTTTTCCCCCACCAGTGGAGCGGCATACATCAGC 2351  
 QY 1737 GGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATATGATGGCATTGT 1796  
 Db 2352 GGGTATGAATTTCCAG-----GACATGGTTTCAGATCCGGAAGAGCGTGGCCCTGTGC 2405  
 QY 1797 CCACAGTTAGATATACACTTTGATGTTTTCAGAGTAGAAGAAATTTATCAATTTGGCT 1856  
 Db 2406 CCGCAGCAGCAGATCCTGTGTTTGACAACTTGACAGTGGCAGAGCACCTTTATTCTACGCC 2465  
 QY 1857 TCAATCAAAGGGATACCGCCAACTAATAATACAAAGAGTGCAGAGGTTTTACTAGAT 1916  
 Db 2466 CAGCTGAAGGCTGTGTCAGTGAAGTGCCTGAAGAGTCAAGCAGATGCTGCACATC 2525  
 QY 1917 TTAGACATGACAGTATCAAGATACCAAGCTAAATAAATTAAGTGGTGGTCAAAAAGA 1976  
 Db 2526 ATCGGCTGGAGGACAAAGTGAACCTCACGAGCCGCTTCTGAGCGGGGCGATAGGCGC 2585  
 QY 1977 AAGCTGTATTAGGAATTCGTTCTTGGGAACCAAGATATCTGCTGTAGATGAACA 2036  
 Db 2586 AAGCTCTCCATCGGCATCCCTCATCGCAGGCTCCAGGTGCTGATCTGAGCAGGCC 2645  
 QY 2037 ACAGCTGAATGAGACCCCTGTTCTGACATATTTGATGAATCTTTTAAAAATACAGAAA 2096  
 Db 2646 ACCTCGGCATGAGCGCCATCTCCAGAGGGGCACTCTGGGATCTTCTTCAGCGGCAGAAA 2705  
 QY 2097 GCCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTGCAGAT 2156  
 Db 2706 AGTGACCGCACCATCGTGTGACCCCACTTCATGGACAGGCTGACCTGTGGAGAC 2765  
 QY 2157 AGGAAAGCTGTGATATCAAGGAATGCTGAATGTGTTGTTCTTCAATGTTCTCTCAA 2216  
 Db 2766 CGCATCGCCATCATGGCCAAAGGGGAGCTGAGTGTGCGGGTCTCTCGCTGTTCTCAAG 2825  
 QY 2217 AGTAAATGGGGATCGGCTACCGCTGACATGATACATAGACAAATATTTGTCACAGAA 2276  
 Db 2826 CAGAAATACGGTCCCGGCTATCAGATGACGCTGGTGAAGAGCGGCATCTGCAACCCGHA 2885  
 QY 2277 TCTCTTTTCTCACTGGTTAAACAAATATACCTGGAGCTACTTTATTAACAAGATGAC 2336  
 Db 2886 GACATCTCCAGCTGGTCCACACACGCTGCCACCGCTGGAGAGCAGCGCTGGG 2945  
 QY 2337 CAACAACCTGTGTATAGCTTGCCTTTCAAGGACATGGACAAATTTTCAGGTTGTTTCT 2396  
 Db 2946 GCGAGCTGTCTTTTCATCTCTCCAGAGAGACGACAGGTTTGAAGGTTCTTTTGT 3005  
 QY 2397 GCCTTAGACAGTCAATCAAAATTTGGGTGTCATTTC---TTATGGTGTTTTCCATGACACT 2453  
 Db 3006 AAACCTGGAGAGAGCAGAAAGAGTGGGCATTCGACGCTTTGGGCGATCCATCACCACC 3065

QY 2454 TTGGAAGAGCTATT 2467  
 Db 3066 ATGGAGGAAGTCTT 3079

RESULT 7

US-09-526-193A-2  
 ; Sequence 2, Application US/09526193A  
 ; Patent No. 6617122  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hayden, Michael R.  
 ; APPLICANT: Brooks-Wilson, Angela R.  
 ; APPLICANT: Pimstone, Simon N.  
 ; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING  
 ; FILE REFERENCE: 50110/002005  
 ; CURRENT APPLICATION NUMBER: US/09/526,193A  
 ; CURRENT FILING DATE: 2000-03-15  
 ; PRIOR APPLICATION NUMBER: 60/124,702  
 ; PRIOR FILING DATE: 1999-03-15  
 ; PRIOR APPLICATION NUMBER: 60/138,048  
 ; PRIOR FILING DATE: 1999-06-08  
 ; PRIOR APPLICATION NUMBER: 60/139,600  
 ; PRIOR FILING DATE: 1999-06-17  
 ; PRIOR APPLICATION NUMBER: 60/151,977  
 ; PRIOR FILING DATE: 1999-09-01  
 ; NUMBER OF SEQ ID NOS: 287  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 7860  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-526-193A-2

Query Match 2.5%; Score 137.6; DB 4; Length 7860;  
 Best Local Similarity 52.7%; Pred. No. 2.9e-22;  
 Matches 349; Conservative 0; Mismatches 304; Indels 9; Gaps 2;

QY 1595 GGAGGCTTTGAGAAATTTGTCATATATGACATATATGAGGTCAGATTAATGCTTACTTGG 1654  
 Db 2813 GTGGCTGTCGATGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 2872  
 QY 1655 CCACAGTGAACAGAGAAAGAGTACATTAATGATGAATATTTTGTGATCTGCTCCACCTTC 1714  
 Db 2873 CCACAAATGAGGGGGAAGACACACCATGTCATCTGACCGGTTGTTCCTCCCGGAC 2932  
 QY 1715 TGATGGTTTGATCTATATATATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGC 1774  
 Db 2933 CTCGGGACCGCTTACCTCTGGGAAAGACANTTGGCTC-----TGAGATGACACCAT 2986  
 QY 1775 AAGAAAAATGATGGCAATTTGTCACAGTTAGATATACACTTTGATGTTTTGACAGTAGA 1834  
 Db 2987 CCGGCAGAACCTGGGGTCTGTCCTCCAGCATAACGTGCTGTTGACATGCTGACTGTGA 3046  
 QY 1835 AGAAAAATTTATCAATTTGGCTTCAATCAAGGGATACAGCAACATATATACAGA 1894  
 Db 3047 AGAACACATCTGGTCTATGTCCTGTTGAAAGGGCTCTCTGAGAGACAGTGAAGCGGA 3106  
 QY 1895 AGTGACAGAGGTTTACTAGATTTAGACATGACAGACTATCAAA---GATAACCAAGCTAA 1951  
 Db 3107 GATGGAGCAGATGGCCCTGGATGTTGTTTGGCATCAAGCAAGCTGAAAGCAAAACAAG 3166  
 QY 1952 AAAATTAAGTGTGTGTCAAAAGAAAGAGTGTGTCATAGGAATGCTGTCTTGGGAACCC 2011  
 Db 3167 CCAGCTGTGAGTGAATGACAGAAAGCTATCTGTGCGCTTGGCCCTTGTGCGGGGATC 3226  
 QY 2012 AAAGATACCTGCTGTAGATGAACCAACAGCTGGAATGACCCCTGTTCTCCGACATATGT 2071  
 Db 3227 TAAGGTTGTCATCTGGATGAACCCACAGCTGTGTGGACCTTACTCCCGCAGGGAAT 3286  
 QY 2072 ATGGAATCTTTTAAANAATACAGAAAGCCAAATCGCGGTGACAGTGTTCAGTACTCATTCAT 2131

Db 3287 ATGGGAGCTGCTGTGAAATACCGAACAGGCGCACCATATTATTCTCTTACACACCAT 3346  
 QY 2132 GGATGAAGCTGACATTTCTTGTCAGATAGGAAAGCTGTGATATCACAAGGAATCTGAAATG 2191  
 Db 3347 GGATGAAGCGGAGCTCTCTGGGGACAGGATTCCTCCATCATCTCCATGGAAGCTGTGCTG 3406  
 QY 2192 TGTGCTTTTCAATGTTCTTCAAAAGTAATAGGGGATCGGCTACCGCTGAGCATGTA 2251  
 Db 3407 TGTGGGCTCTCTCCCTGTTTCTGAAGAACAGCTGGGAAACAGGCTACTACCTGACCTTGT 3466  
 QY 2252 CA 2253  
 Db 3467 CA 3468

RESULT 8

US-08-858-207A-88  
 ; Sequence 88, Application US/08858207A  
 ; Patent No. 6348328  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Black, Michael  
 ; APPLICANT: Hodgson, John  
 ; APPLICANT: Knowles, David  
 ; APPLICANT: Nicholas, Richard  
 ; APPLICANT: Stodola, Robert  
 ; TITLE OF INVENTION: No. 6348328e1 Compounds  
 ; NUMBER OF SEQUENCES: 552  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SmithKline Beecham Corporation  
 ; STREET: 709 Swedeland Road  
 ; CITY: King of Prussia  
 ; STATE: PA  
 ; COUNTRY: USA  
 ; ZIP: 19406-0939  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/858,207A  
 ; FILING DATE: 09-MAY-1997  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/017670  
 ; FILING DATE: 14-MAY-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Gimmi, Edward R.  
 ; REGISTRATION NUMBER: 38,891  
 ; REFERENCE/DOCKET NUMBER: P50475  
 ; TELEPHONE: 610-270-4478  
 ; TELEFAX: 610-270-5090  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 88:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1064 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 US-08-858-207A-88

Query Match 1.9%; Score 103.4; DB 4; Length 1064;  
 Best Local Similarity 47.9%; Pred. No. 1.2e-14;  
 Matches 328; Conservative 0; Mismatches 356; Indels 1; Gaps 1;  
 QY 4006 AATATGATGACAGAAAGATTTTCTTCTTCAAGAAAGTAAGAAAGTGGCAACTAAAT 4065  
 Db 143 AAAAGGAGGCAAAAGTCTTTTATCCATCCCGAAAAACAAATTTTGAAGCGGTTAAG 202  
 QY 4066 ACATCTCTTCTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGGTGCTG 4125  
 Db 203 ATTATCTGTTGAGGTTCCAAAGGACAAATTTTAGTTTATTGTTGTTATGGGCGAG 262





TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 309-8504  
TELEFAX: (301) 309-8512  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5625 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-08-961-527-1

Query Match 1.7%; Score 92.2; DB 4; Length 5625;  
Best Local Similarity 46.2%; Pred. No. 9.9e-152;  
Matches 307; Conservative 0; Mismatches 358; Indels 0; Gaps 0;

QY	4026	TTTCTTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTTCTGTGTGAAA	4085
Db	2132	TTTATTCATCTCTGAAAGCAGACCTTTGAAGCGGTCAAGGATTTGACCTTTGAGGTTCCA	2191
QY	4086	AAAGAGAGATCTTAGGACTATGGTCCAAATGGTCTGGCAAAAGCACAAATTATTAAT	4145
Db	2192	AAAGGCGAGATTTAGGATTTATCGGGCAAAATGGTCTGGGAAGTCGACAAACCAATTAAA	2251
QY	4146	ATTCGTGGTGGTATTCGAACCACTTCAGGCCAGGTATTTTAGGAGATTTATCTTCA	4205
Db	2252	ATGCTGACAGGATTTTGAACCAACATCTGTTTTTGTGCGATTAACGGCAAGATTCCC	2311
QY	4206	GAGACAAGTGAAGATGAATTCATCTGAAGTGTATGGTTACTGCTCTCAGATAAACCTT	4265
Db	2312	CAGGACAATCGGCAAGATTTATGTCAAAGATAATGGCGTAGTCTTTGGCAACAGCCACG	2371
QY	4266	TTGTGCCAGATCTACTATGTCAGGAACATTTTGAATTTATGGAGCTGTCAAAGGAATG	4325
Db	2372	CTATGGTGGATTTGGCTCTGCAAGAGACCTACACTGCTTAAAGAGATTTATGATGTG	2431
QY	4326	AGTGAAGTGACATGAAGAAGCTATAAGTCGAATPACACATGCACTTGATTTAAAGAA	4385
Db	2432	CCAGACTCGCTCTTTCATAAGCGTATGGACTTTTGAATGAAGTCTTGATTTGAAGGAC	2491
QY	4386	CATCTTCAGAGACTGTAAGAACTACTCGCAGCAATCAACGAAAGTTGTGTTGCT	4445
Db	2492	TTTATCAAGATCCCGTGGGACTCTTTCAGTGGACAAACGATCGCGCGGATATTCG	2551
QY	4446	CTAAGTATCTAGGAATCCTCAGATTAATTTGTAGATGAACCATCTACAGGTATGGAT	4505
Db	2552	GCCTCCTGCTCCCAATCCCAAGGTTCTTTTTTTAGATGAGCCGACCATTTGGTTGGAC	2611
QY	4506	CCCAAGCCCAACAGCACATGTGGCGAGCAATTCGAATGCAATTTAAACACAGAAAGCG	4565
Db	2612	GTTCGGTTAAGGATAATATTCGTCGGGCAATTAATCTCAGATCAATCAAGAGAAAGAACT	2671
QY	4566	GCTGCTATTCTGACCACTCACTATATGAGAGAGGAGCGGTGCTGTGATCGAGTAGCT	4625
Db	2672	ACCATCTTTTGACCACTCACGATTTAGTGATATTGAGCAACTTTGTGATCGGATTTTC	2731
QY	4626	ATCATGGTCTGCGGAGTTAGATGTATCGGAACAGTACACATCTAAAGATTAATTT	4685
Db	2732	ATGATTGACAAGGGGCAAGAGATTTTGTATGGAACCGGTGAGCCAACTCAAGGAGACCTTT	2791
QY	4686	GGAAA 4690	
Db	2792	GGTAA 2796	

RESULT 11

US-09-621-976-2582  
Sequence 2582, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.

FILE REFERENCE: GENSET.054PR2

CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 2582  
LENGTH: 450  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 69..308  
US-09-621-976-2582

Query Match 1.5%; Score 83.8; DB 4; Length 450;  
Best Local Similarity 60.2%; Pred. No. 3.2e-10;  
Matches 139; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY	4796	AACTTTTCTTCTATTTTGGCTTATAAAATTTCTTAAGGAGATGTTCACTCCCTTTTCA	4855
Db	74	AAATATTTCTCTTTAATGGGTATTAAGTACTGTGGAGATGTCACCCCTCTATCTCG	133
QY	4856	ATCTTTTAAAGCTGGAAGAAGCTAAACATGCTTTTGGCATTTGAAGAATATAGCTTTTC	4915
Db	134	GGCTTTTCAAGTTAGAGCGGATGAACAGACCTTCAACCTGGAGGAATACAGCCTCTC	193
QY	4916	TCAGCAACATTTGGAACAGGTTTTTGTAGACTCATAAGAACAGAGAGGAAGATAA	4975
Db	194	TCAGGCTACCTTGGAGCAGGTATTTCTAGAACTCTGTAAGAGCAGGAGCTGGGAATGT	253
QY	4976	TAGTGTGGAACCTTTAAACAGCACACTTTGGTGGGAACGAACACACAGAAGA	5026
Db	254	TGATGATAAAATTGATACACAGCTTGAATGGAACTTCTCCACACAGAAGA	304

RESULT 12

US-09-621-976-2581  
Sequence 2581, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 2581  
LENGTH: 463  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 110..328  
US-09-621-976-2581

Query Match 1.5%; Score 83.4; DB 4; Length 463;  
Best Local Similarity 60.3%; Pred. No. 4.1e-10;  
Matches 138; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY	4798	GTCTTTCTTCTATTTTGGCTTATAAAATTTCTTAAGGAAGATGTTCACTCCCTTTCA	4857
Db	96	GATATTTCTCTTTAATGGCGTATAAGTTACCTGTGGAGATGTCCACCTCTATCTCGC	155
QY	4858	CTTTTAAAGCTGGAAGAAGCTAAACATGCTTTTGGCATTTGAAGAATATAGCTTTTCTC	4917
Db	156	CTTTTCAAGTTAGAGCGGATGAACAGACCTTCAACCTGGAGGAATACAGCTCTCTC	215
QY	4918	AAGCAACATTTGGAACAGGTTTTTGTAGAACTACTAAGAACAGAGGAGGAGATAA	4977
Db	216	AGGCTACCTTGGAGCAGGTATTTCTTAGAACTCTGTAAGAGCAGGAGCTGGGAATGTTG	275

QY 4978 GTTGTGGACCTTAAACAGCACACTTTGGTGGGAGAGAACACACAGAGA 5026  
 Db 276 ATGATAAAATGATACACAGTTGAATGGAACCTTCTCCACAGGAAGA 324

RESULT 13  
 US-09-621-976-2583  
 ; Sequence 2583, Application US/09621976  
 ; Patent No. 6639053  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Jobert, S.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 2583  
 ; LENGTH: 478  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: 110..328  
 US-09-621-976-2583

Query Match 1.5%; Score 83.4; DB 4; Length 478;  
 Best Local Similarity 60.3%; Pred. No. 4.1e-10;  
 Matches 138; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
 QY 4798 GTTTTCTTCTTATTTTGGCTTATATAAATTCCTAAGGAGATGTTTCAGTCCCTTTTCACAAT 4857  
 Db 96 GATATTCCTTTTAAATGCGGTATAGTTACCTGTGGAGATGTCACCCCTCTATCTCGGC 155  
 QY 4858 CTTTCTTTTAACTGGAAGCAAGCAATGCTTTTGGCAATGAGATATAGCTTTTCTC 4917  
 Db 156 CTTTCTTTTCAAGTTAGGCGATGAACACAGCTTCAACCTGGAGATACAGCTCTCTC 215  
 QY 4918 AAGCAACTTGAACAGCTTTTGTAGAACTCCTAAGAAACAGAGAGAGAGATATA 4977  
 Db 216 AGGCTACCTTGGAGCAGGTATCTTAGAACTCTGTAAAGACAGAGCTGGGAATCTTG 275  
 QY 4978 GTTGTGGAACCTTAAACAGCACACTTTGGTGGGAGAGAACACACAGAGA 5026  
 Db 276 ATGATAAAATGATACACAGTTGAATGGAACCTTCTCCACAGGAAGA 324

RESULT 14  
 US-09-833-381-1080  
 ; Sequence 1080, Application US/09833381  
 ; Patent No. 6672186  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Robison, Keith E.  
 ; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
 ; FILE REFERENCE: 5800-119  
 ; CURRENT APPLICATION NUMBER: US/09/833,381  
 ; CURRENT FILING DATE: 2001-04-11  
 ; PRIOR APPLICATION NUMBER: 09/516,448  
 ; PRIOR FILING DATE: 2000-02-29  
 ; NUMBER OF SEQ ID NOS: 2050  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1080  
 ; LENGTH: 465  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)..(465)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-833-381-1080

Query Match 1.5%; Score 80.8; DB 4; Length 465;  
 Best Local Similarity 57.2%; Pred. No. 1.6e-09;  
 Matches 207; Conservative 0; Mismatches 147; Indels 8; Gaps 3;  
 QY 3796 ATGAGAAAAATATGAGGAGATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTCAA 3855  
 Db 110 ATATAATGAACTGAGGAGAAACATATGAGAAAGGATCTGTGTTTCAGAAATTTCTCAA 169  
 QY 3856 CGAAGTCTTAAATAGGAAGCTTCCAGAACCCAGACAAATGAGGATGAAGATGAAGATG 3915  
 Db 170 -----GAAGCAACGCTATTTTCCAAACCCAGAGAGCGCTGAAGGAGAGGAGAGATA 223  
 QY 3916 TCAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGTTGCCAGTGTCTGAGGAGAAAC 3975  
 Db 224 TCCAGATGGAAGAAATGAGAAACAGTCAATGCTATGGCTGCGAGACTTTGATGAGACAC 283  
 QY 3976 CATCCATTATGGTCAGCAATTTGCATAAA-GAATATGATGACAAAGAAATTTCTCTT 4034  
 Db 284 CGTCATCATTTGCCAGCTGTCTACCGGAAGGAATATGACGGCAAAAAGAAAAATTTGCTTT 343  
 QY 4035 TCAAGAAAGTAAAGAAAGTGGCAACTAATACATCTCTTTCTGTGTAAGAAAGGAG-A 4093  
 Db 344 TCTAAAGGAAGAAACAAATGCCCACAAAGAAATGCTCTTTTGTGTTAAAAAGGTGAA 403  
 QY 4094 GATCTTAGGACTATTGGGTCCAAATGGTGTCTGCGCAAAAGCACAAATTTAATTTCTG 4153  
 Db 404 GTTATAGGACTGTTAGGACACAAATGGAGCTGTTAAAGTACAACTATTAAAGATGATAAC 463  
 QY 4154 TG 4155  
 Db 464 TG 465

RESULT 15  
 US-08-961-527-264  
 ; Sequence 264, Application US/08961527  
 ; Patent No. 6420135  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Charles Kunsch  
 ; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
 ; NUMBER OF SEQUENCES: 391  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
 ; COMPUTER: HP Vectra 486/33  
 ; OPERATING SYSTEM: MSDOS version 6.2  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/961,527  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Brookes, A. Anders  
 ; REGISTRATION NUMBER: 36,373  
 ; REFERENCE/DOCKET NUMBER: PB340P1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301) 309-8504  
 ; TELEFAX: (301) 309-8512  
 ; INFORMATION FOR SEQ ID NO: 264:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 795 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double

```

; TOPOLOGY: linear
US-08-961-527-264

Query Match      1.5%; Score 80.6; DB 4; Length 795;
Best Local Similarity 48.3%; Pred. No. 2.3e-09;
Matches 320; Conservative 0; Mismatches 334; Indels 9; Gaps 3;

QY 4048 AGAAGTGGCAACTAAATACATCTCTTCTGTGTGAAAGAGAGAGATCTTAGGACTAT 4107
Db      |||||
3 ATAAAGAGGCTTGAGTAATCTCAATTTGCAGATTGAAATGGAGAGATTATGGGCTTGA 62

QY 4108 TGGGTCCAAATGGTCTGCGCAAGACCAATATATATTTCTGTTGGTTCATATTCGAC 4167
Db      |||||
63 TTGGTCATAATGGGGCTGGAAATCGACCCTATAAATCCCTAGTCAGTATCAATTTCCAC 122

QY 4168 CAATCTCAGGCGAGGTATTTTAGGAGATTATCTTCAGAGACAAGTGAAGATGATGATT 4227
Db      |||||
123 CCAGCAGTGGTCTGATTTTGTAGACGGTACGGAGTTATCGGAAAATCGCTTGGCTATTA 182

QY 4228 CACTGAAGTATGGGTTACTGTCTCTAGATAAACCCCTTTGTGGCCAGATACATTCG 4287
Db      |||||
183 AACGAAAG---ATTGGCTACGTAGCAGACTCGCCTGACTTATTTTACGCTTAACGGCCA 239

QY 4288 AGGAACATTTGAAATTTATGAGCTGTCAAGGAATGAGTCAAGTGAATGAAGAAG 4347
Db      |||||
240 ATGATTTTGGGATTTGATCGCCTCATCTATGATCTGAGTAGATCTGACTTGGAGGCTA 299

QY 4348 TCATAAGTCGAATAACACATGCACTTGATTTTAAAGAACATCTTCAGAACTGTAAAGA 4407
Db      |||||
300 GTCTAGCTAGGCTATTGAACGCTTTTGTGATTTTGTGAAAATCGCTATCAGGTTATTGAAA 359

QY 4408 AACTACCTGCGAGGATCAACGAAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCCTC 4467
Db      |||||
360 CTCCTTCTCAGGAATGCGTCAGAAAGTCTTTGTCTATCGGAGCACTCTTGTCTGATCCCG 419

QY 4468 AGATTACTTTGCTAGATGAACCATCTACAGTATGGATCCCAAAGCCAAACAGACATGT 4527
Db      |||||
420 ATATTTGGGCTCTTGGATGAACCCCTGACTGGTTTGGATCCCGAGGCTGCCCTTGA---TT 476

QY 4528 GCGGAGCAATTCGAACTCAATTTAAACAGAAAGCGGCTGCTATTCTGACCACTCACT 4587
Db      |||||
477 TGAACAGATGATGAAGGACATGCAAAAAGGGAGAGACAGTCTTGTTTTCAACTCATG 536

QY 4588 ATATGAGAGGCGCAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGCGAGTTAA 4647
Db      |||||
537 TCCTAGAGGTGGCAGAGCAAGTCTGTGATCGGATTCGATTTTGAAGGAGGCGCATTTGA 596

QY 4648 GATGTATCGGAACAGTACACATCTAAAGAGTAAATTTGGAAGGCTACTTTTGGAAA 4707
Db      |||||
597 TTTATTGGTGTAGTGTAGGACTT---GAGAAAAGATTACCCAGACCGATCTTTGGAAA 653

QY 4708 TTA 4710
Db      ||
654 GTA 656

```

Search completed: April 12, 2004, 07:56:05  
Job time : 260 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2004, 15:37:07 ; Search time 1244 Seconds

(Without alignments)  
16510.653 Million cell updates/sec

Title: US-10-090-458-4

Perfect score: 5475

Sequence: 1 GCGTCGCGCCCTCGCACAGT.....cataaagcaatggaagtt 5475

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2475585 seqs, 1875730760 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4951170

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/FCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5475	100.0	5475	13	US-10-090-458-4
2	5181	94.6	6525	14	US-10-005-338B-1
3	4664	85.2	5262	9	US-09-971-121-5
4	4572	83.5	4929	9	US-09-971-121-1
5	4408	80.5	4785	9	US-09-971-121-3
6	2826	51.6	5463	13	US-10-090-458-1
7	2739	50.0	3268	13	US-10-094-749-984
8	2713	49.6	4917	13	US-10-090-458-3
9	1991	36.4	3347	15	US-10-108-260A-160
10	1497	27.3	1548	10	US-09-822-846-97
11	1343	24.5	3928	10	US-09-822-846-98
12	744	13.6	1818	14	US-10-204-887-32
13	743	13.6	974	9	US-09-833-381-1085
14	742	13.6	1943	15	US-10-108-260A-1424
15	459	8.4	465	12	US-10-085-783A-44227

## ALIGNMENTS

### RESULT 1

```

US-10-090-458-4
; Sequence 4, Application US/10090458
; Publication No. US20020123107A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hongyun
; APPLICANT: Kilinski, Ligia
; APPLICANT: Le Bihan, Stephane
; TITLE OF INVENTION: NOVEL ABCA5 TRANSPORTER AND USES THEREOF
; FILE REFERENCE: 100103.403
; CURRENT APPLICATION NUMBER: US/10/090,458
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 5475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 25
; OTHER INFORMATION: n = A,T,C or G
US-10-090-458-4

```

```

Query Match      100.0%; Score 5475; DB 13; Length 5475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTCGCGCCCTCGCACAGATCCNAGCTGTCACCGCAGTCAGTCAACAGACTCGAGC 60
DB 1 GCGTCGCGCCCTCGCACAGATCCNAGCTGTCACCGCAGTCAGTCAACAGACTCGAGC 60
QY 61 GCGTCGCGCCCTCGCACAGCTCTCGGCTCGGCGCCCTGAGGTTTATTGAGAAAACATGTCCA 120
DB 61 GCGTCGCGCCCTCGCACAGCTCTCGGCTCGGCGCCCTGAGGTTTATTGAGAAAACATGTCCA 120
QY 121 CTGCAATTAGGAGGTAGGAGTTTGAGACACAGACACACTTCTACTGAGAAATTACT 180

```

```

Sequence 44227, A
Sequence 16105, A
Sequence 16105, A
Sequence 28509, A
Sequence 28509, A
Sequence 10804, A
Sequence 337, App
Sequence 3406, App
Sequence 1033, App
Sequence 333, App
Sequence 4362, App
Sequence 1883, App
Sequence 4371, App
Sequence 22, Appl
Sequence 43535, A
Sequence 43535, A
Sequence 5, Appl
Sequence 973, App
Sequence 115303,
Sequence 115304,
Sequence 971, App
Sequence 11289, A
Sequence 27989, A
Sequence 16501, A
Sequence 2801, App
Sequence 9786, App
Sequence 15066, A
Sequence 15066, A
Sequence 10033, A

```

```

15 US-10-242-535A-44227
12 US-10-085-783A-16105
456 12 US-10-242-535A-16105
456 12 US-10-085-783A-28509
473 12 US-10-242-535A-28509
473 15 US-09-867-701-10804
477 9 US-09-862-436-337
1346 9 US-09-880-107-3406
1346 9 US-09-954-531-1033
1346 14 US-10-171-581-333
476 9 US-09-867-701-4362
469 9 US-09-880-107-1893
412 9 US-09-867-701-4371
431 15 US-10-264-237-22
252 12 US-10-085-783A-43535
252 15 US-10-242-535A-43535
2079 9 US-09-767-870-5
2079 14 US-10-242-568-5
7838 15 US-10-242-355-973
709 15 US-10-027-632-115303
709 15 US-10-027-632-115304
1478 15 US-10-242-355-971
474 9 US-09-864-761-11289
180 9 US-09-864-761-27989
147 14 US-10-029-386-16501
566 14 US-10-029-386-2801
315 9 US-09-783-590-9786
412 12 US-10-085-783A-15066
412 15 US-10-242-535A-15066
60 10 US-09-908-975-10033

```

```

16 459 8.4
17 450 8.2
18 456 8.2
19 450 8.2
20 450 8.2
21 448 8.2
22 376 6.9
23 376 6.9
24 376 6.9
25 376 6.9
26 331 6.0
27 324 5.9
28 299 5.5
29 284 5.2
30 252 4.6
31 252 4.6
32 243 4.4
33 243 4.4
34 213 3.9
35 207 3.8
36 207 3.8
37 193 3.5
38 178 3.3
39 164 3.0
40 144 2.6
41 144 2.6
42 128 2.3
43 121 2.2
44 121 2.2
45 39 0.7

```

Db 121 CTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACCAGAACACTTCTACTAGAAATTACT 180  
Qy 181 TAAATTAAATGAGAACCAAAAAGAGTAGTGTTCCGGAATTCCTTTTCCACTATTTTTT 240  
Db 181 TAAATTAAATGAGAACCAAAAAGAGTAGTGTTCCGGAATTCCTTTTCCACTATTTTTT 240  
Qy 241 TATTTTGGTAAATATTAATTAAGCATGATGCATCCAAATTAAGAAATATGAAGAGTGCTTA 300  
Db 241 TATTTTGGTAAATATTAATTAAGCATGATGCATCCAAATTAAGAAATATGAAGAGTGCTTA 300  
Qy 301 ATATAGAACTCAATCCTATGACCAAGTTTACTTTCTPAATCTAAATCTTTGGATATATC 360  
Db 301 ATATAGAACTCAATCCTATGACCAAGTTTACTTTCTPAATCTAAATCTTTGGATATATC 360  
Qy 361 CAGTGACTAATATACAGAGAGCATCATGACAGAAAGTGTCTACTGATCATCTACCTGATG 420  
Db 361 CAGTGACTAATATACAGAGAGCATCATGACAGAAAGTGTCTACTGATCATCTACCTGATG 420  
Qy 421 TCATAAATPACTGAAGAATPATACAAATGAAAAGAAATGTTAAACATCCAGTCTCTTAAGC 480  
Db 421 TCATAAATPACTGAAGAATPATACAAATGAAAAGAAATGTTAAACATCCAGTCTCTTAAGC 480  
Qy 481 CGAGCAACTTTGTAGGTGTGTTTTCAAGACTCCATGTCCTATGAACTTCGTTTTTTTC 540  
Db 481 CGAGCAACTTTGTAGGTGTGTTTTCAAGACTCCATGTCCTATGAACTTCGTTTTTTTC 540  
Qy 541 CTGATATGATCCAGTATCTTCTATTTATATATGATTCAGAGCTGGCTGTTCAAAATCAT 600  
Db 541 CTGATATGATCCAGTATCTTCTATTTATATATGATTCAGAGCTGGCTGTTCAAAATCAT 600  
Qy 601 GTGAGGCTGCTCAGTACTGTCCTCAGGTTTCAAGTTTTACAGCATTCATAGATGCTG 660  
Db 601 GTGAGGCTGCTCAGTACTGTCCTCAGGTTTCAAGTTTTACAGCATTCATAGATGCTG 660  
Qy 661 CCATATATACAGTGAAGACCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACTAAAGCTG 720  
Db 661 CCATATATACAGTGAAGACCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACTAAAGCTG 720  
Qy 721 TTATATGGGAGAACTGCTGTTAGAAATAGATACCTTTCCCGAGGAGTAAATTTAA 780  
Db 721 TTATATGGGAGAACTGCTGTTAGAAATAGATACCTTTCCCGAGGAGTAAATTTAA 780  
Qy 781 TATACCTAGTTATAGCATTTTCACTTTGGATACTTTTTGGCAATTCATATGCTAGCAG 840  
Db 781 TATACCTAGTTATAGCATTTTCACTTTGGATACTTTTTGGCAATTCATATGCTAGCAG 840  
Qy 841 AAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACTGCCTTTT 900  
Db 841 AAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACTGCCTTTT 900  
Qy 901 GGCTTTCCTGGTCTCTATATACAGTTTAAATTTTCTATGTCCTCTTATGSCAG 960  
Db 901 GGCTTTCCTGGTCTCTATATACAGTTTAAATTTTCTATGTCCTCTTATGSCAG 960  
Qy 961 TCATTGGACAGCTCTTTGTTTATTTCTCTCAAGTAGCAGCATTTGTGATATTTCTGCTT 1020  
Db 961 TCATTGGACAGCTCTTTGTTTATTTCTCTCAAGTAGCAGCATTTGTGATATTTCTGCTT 1020  
Qy 1021 TTTTCTTTATGGATATCATCTGATTTTTTTGCTTTTAAAGTCTGACACCTCTTTTAAA 1080  
Db 1021 TTTTCTTTATGGATATCATCTGATTTTTTTGCTTTTAAAGTCTGACACCTCTTTTAAA 1080  
Qy 1081 AATCAAAACATGTGGGAATAGTTGAATTTTTTTGTTTACTGTCCTTTTGGATTTATTTGGCC 1140  
Db 1081 AATCAAAACATGTGGGAATAGTTGAATTTTTTTGTTTACTGTCCTTTTGGATTTATTTGGCC 1140  
Qy 1141 TTATGATTAATCCTCATAGAAAGTTTCCAAATCGTTAGTGTGGCTTTTCAGTCCCTTCT 1200  
Db 1141 TTATGATTAATCCTCATAGAAAGTTTCCAAATCGTTAGTGTGGCTTTTCAGTCCCTTCT 1200  
Qy 1201 GTCACTGTACTTTGTGATGCTGATTTGACAGGTCATGACATTTAGAGATTTTAAATGAAG 1260  
Db 1201 GTCACTGTACTTTGTGATGCTGATTTGACAGGTCATGACATTTAGAGATTTTAAATGAAG 1260

Qy 1261 GTGCTTCATTTTCAAAATTTGACTGCGGCCCATATCTCTCTAATTAATTAATTAATTAATTCATGC 1320  
Db 1261 GTGCTTCATTTTCAAAATTTGACTGCGGCCCATATCTCTCTAATTAATTAATTAATTAATTCATGC 1320  
Qy 1321 TCACACTTAATAGTATATTTCTATGTCCTCTGGCTGTCTATCTTGATCAAGTCATTCCAG 1380  
Db 1321 TCACACTTAATAGTATATTTCTATGTCCTCTGGCTGTCTATCTTGATCAAGTCATTCCAG 1380  
Qy 1381 GGGAAATTTGGCTTAGCGGATCATCTTTATATTTTCTGAAAGCCCTTCATATTTGGTCAAGA 1440  
Db 1381 GGGAAATTTGGCTTAGCGGATCATCTTTATATTTTCTGAAAGCCCTTCATATTTGGTCAAGA 1440  
Qy 1441 GCAAAAGAAATTTATCAGGAGTTTATCAGAGGCAATGTTAAATGGAATATTTAGTTTTAGTG 1500  
Db 1441 GCAAAAGAAATTTATCAGGAGTTTATCAGAGGCAATGTTAAATGGAATATTTAGTTTTAGTG 1500  
Qy 1501 AAATATTGAGCCAGTTTCTCAGAAATTTGTAGGAAAAGAGCCATAGAAATTTAGTGTA 1560  
Db 1501 AAATATTGAGCCAGTTTCTCAGAAATTTGTAGGAAAAGAGCCATAGAAATTTAGTGTA 1560  
Qy 1561 TTCAAGACATACAGAAAGAGGTTGAAATGTCGAGGCTTTGAGAAATTTGTCTATTG 1620  
Db 1561 TTCAAGACATACAGAAAGAGGTTGAAATGTCGAGGCTTTGAGAAATTTGTCTATTG 1620  
Qy 1621 ACATATATGAGGCTCAGATTTACTGCTTTACTTTGGCCACAGTGGAAACAGAAAGATACAT 1680  
Db 1621 ACATATATGAGGCTCAGATTTACTGCTTTACTTTGGCCACAGTGGAAACAGAAAGATACAT 1680  
Qy 1681 TGATGAATATCTTTTGTGACTCTGCCACCTCTCTGATGGGTTTGCATCTATATATGAC 1740  
Db 1681 TGATGAATATCTTTTGTGACTCTGCCACCTCTCTGATGGGTTTGCATCTATATATGAC 1740  
Qy 1741 ACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATATGATTTGGCATTTCTCCAC 1800  
Db 1741 ACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATATGATTTGGCATTTCTCCAC 1800  
Qy 1801 AGTTAGATATACATTTGATGTTTTGACAGTAGAAGAAATTTATCAATTTTGGCTTCAA 1860  
Db 1801 AGTTAGATATACATTTGATGTTTTGACAGTAGAAGAAATTTATCAATTTTGGCTTCAA 1860  
Qy 1861 TCAAGGGATACCGCCCAACATAATAACAAGAGTGCAGAAAGTTTACTAGATTTAG 1920  
Db 1861 TCAAGGGATACCGCCCAACATAATAACAAGAGTGCAGAAAGTTTACTAGATTTAG 1920  
Qy 1921 ACATGACAGCTATCAAGATTAACAGCTTAAATAATTAAGTGGTGTCAAAAAGAAAGC 1980  
Db 1921 ACATGACAGCTATCAAGATTAACAGCTTAAATAATTAAGTGGTGTCAAAAAGAAAGC 1980  
Qy 1981 TGTCAATTAGGAATTCGTCTTTCTGGGAAACCAAGATATCTGCTAGATGAACCAACAG 2040  
Db 1981 TGTCAATTAGGAATTCGTCTTTCTGGGAAACCAAGATATCTGCTAGATGAACCAACAG 2040  
Qy 2041 CTGGAATGGAACCCCTGTTCTGACATATTTGATGGAATCTTTTAAATACAGAAAGCCA 2100  
Db 2041 CTGGAATGGAACCCCTGTTCTGACATATTTGATGGAATCTTTTAAATACAGAAAGCCA 2100  
Qy 2101 ATCGGTGACAGTGTTCAGTACTCTATTTTCATGGAATGAGCTGACATTTCTGAGATAGA 2160  
Db 2101 ATCGGTGACAGTGTTCAGTACTCTATTTTCATGGAATGAGCTGACATTTCTGAGATAGA 2160  
Qy 2161 AAGCTGTGATATCAAGAAATGCTGAAATGTTGTTGTTCTCAATGTTCTCATAAAGTA 2220  
Db 2161 AAGCTGTGATATCAAGAAATGCTGAAATGTTGTTGTTCTCAATGTTCTCATAAAGTA 2220  
Qy 2221 AATGGGGATCGGCTACCGCTGAGCATGTACATAGCAAAATATTTGTCGCCACAGATCTC 2280  
Db 2221 AATGGGGATCGGCTACCGCTGAGCATGTACATAGCAAAATATTTGTCGCCACAGATCTC 2280  
Qy 2281 TTTCTTCACTGGTAAACACATATACCTGGAGCTACTTTTATTAACAAGATGACCAAC 2340  
Db 2281 TTTCTTCACTGGTAAACACATATACCTGGAGCTACTTTTATTAACAAGATGACCAAC 2340

QY 2341 AACTTGTGTATAGCTTGCCTTTCAAGGACATGGCAAAATTTTCAGGTGTTGTTTCTGCCC 2400  
Db AACTTGTGTATAGCTTGCCTTTCAAGGACATGGCAAAATTTTCAGGTGTTGTTTCTGCCC 2400  
QY 2401 TAGACAGTCATTCAAATTTGGGTGTCAATTTCTTATGGTGTTCATGACGACTTTGGAAG 2460  
Db TAGACAGTCATTCAAATTTGGGTGTCAATTTCTTATGGTGTTCATGACGACTTTGGAAG 2460  
QY 2461 ACGTATTTTAAAGCTAGAGTGTGAAGCAGAAATGGACCAAGCAGATATAGTGTATTTA 2520  
Db ACGTATTTTAAAGCTAGAGTGTGAAGCAGAAATGGACCAAGCAGATATAGTGTATTTA 2520  
QY 2521 CTCAGCACCCTCGAGGAAGAAATGGATTCAAATCTTTTGTGAAATGGAACAGAGCT 2580  
Db CTCAGCACCCTCGAGGAAGAAATGGATTCAAATCTTTTGTGAAATGGAACAGAGCT 2580  
QY 2581 TACTTATTTCTTGAACCAAGGCTTCTAGTGAGCACCACATGAGCTTTGGAAACAAC 2640  
Db TACTTATTTCTTGAACCAAGGCTTCTAGTGAGCACCACATGAGCTTTGGAAACAAC 2640  
QY 2641 AGATGTATACAAATAGCAAAAGTTTCATTTCTTTTACCTTGAAACGTGAAAGTAAATCAGTGA 2700  
Db AGATGTATACAAATAGCAAAAGTTTCATTTCTTTTACCTTGAAACGTGAAAGTAAATCAGTGA 2700  
QY 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTGTTTC 2760  
Db GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTGTTTC 2760  
QY 2761 ATCACTCTTTTAAATATGCTGTGTTCCCATCAAACTTGTTCAGACCTATATTTTCTAA 2820  
Db ATCACTCTTTTAAATATGCTGTGTTCCCATCAAACTTGTTCAGACCTATATTTTCTAA 2820  
QY 2821 AACCTGGAGACAAACCAATAAATACAAACAAAGTCTGCTTCTTCAAAATCTGCTGACT 2880  
Db AACCTGGAGACAAACCAATAAATACAAACAAAGTCTGCTTCTTCAAAATCTGCTGACT 2880  
QY 2881 CAGATATCAGTGATCTTATAGCTTTTTCACAGCCAGACATATGCTGACGATGTTA 2940  
Db CAGATATCAGTGATCTTATAGCTTTTTCACAGCCAGACATATGCTGACGATGTTA 2940  
QY 2941 ATGACAGTGACTATGTATCCGTGGCTCCCATAGTGGGCTTTAAATGTGATGACTTCAG 3000  
Db ATGACAGTGACTATGTATCCGTGGCTCCCATAGTGGGCTTTAAATGTGATGACTTCAG 3000  
QY 3001 AAAGGACTATGTTTTCAGCTGTTTTCACAGCTACTATGTTTATCTTACCTATAT 3060  
Db AAAGGACTATGTTTTCAGCTGTTTTCACAGCTACTATGTTTATCTTACCTATAT 3060  
QY 3061 TAGTGAATATCAATAGTAACCTACTATCTTATCAATTAATGTGACTGAAACCAATCCAGA 3120  
Db TAGTGAATATCAATAGTAACCTACTATCTTATCAATTAATGTGACTGAAACCAATCCAGA 3120  
QY 3121 TCTGAGGTACCCCATCTTTTCAGAAATTAATGATATAGTGTGTTTAAATTTGAGCTGTATT 3180  
Db TCTGAGGTACCCCATCTTTTCAGAAATTAATGATATAGTGTGTTTAAATTTGAGCTGTATT 3180  
QY 3181 TTCAGCAGCTTTGCTGGAATCATGTTACTGCAATGCGACCTTACTTTTGCCATGGAAA 3240  
Db TTCAGCAGCTTTGCTGGAATCATGTTACTGCAATGCGACCTTACTTTTGCCATGGAAA 3240  
QY 3241 ATGCAGAGAATCATAAGATCAAGCTTATATCTCAACTTAACTTTTCAAGTCTTTTGGCAT 3300  
Db ATGCAGAGAATCATAAGATCAAGCTTATATCTCAACTTAACTTTTCAAGTCTTTTGGCAT 3300  
QY 3301 CTGCATATTTGGAATGGACAAGCTGTTGTGATATCCCTTATTTTATCATCTTATTTT 3360  
Db CTGCATATTTGGAATGGACAAGCTGTTGTGATATCCCTTATTTTATCATCTTATTTT 3360  
QY 3361 TGATGCTAGGAAGCTTATGTCATTTTCAATATGGAATATATTTTATGCTGAAAGTTCC 3420  
Db TGATGCTAGGAAGCTTATGTCATTTTCAATATGGAATATATTTTATGCTGAAAGTTCC 3420  
QY 3421 TTGCTGTGGTTTTGTGCTTATGGTATGTTTCCATCAGTATCTGTTTCACTTATATTG 3480

Db 3421 TTGCTGTGGTTTTGTGCTTATGGTATGTTTCCATCAGTATCTGTTTCACTTATATTG 3480  
QY 3481 CTTCTTTACCTTTAAGAAAAATTTTAAATPACAAAGAAATTTTGGTCATTTATCTATCTG 3540  
Db CTTCTTTACCTTTAAGAAAAATTTTAAATPACAAAGAAATTTTGGTCATTTATCTATCTG 3540  
QY 3541 TGGCAGCGTGGCTGTATTTGCAATCAGTGAATTAACCTTTCTTTATGGGATACACAATG 3600  
Db TGGCAGCGTGGCTGTATTTGCAATCAGTGAATTAACCTTTCTTTATGGGATACACAATG 3600  
QY 3601 CAACTATTTCTTCAATATGCTTTTGTATCATCATTCATTCATCTTACCTTCTAGGTGCTC 3660  
Db CAACTATTTCTTCAATATGCTTTTGTATCATCATTCATTCATCTTACCTTCTAGGTGCTC 3660  
QY 3661 TGAATTTCTTTTCAATAGATTTCTTGGAAAGATGTAGAAAAAATGTGGACACCTATAATC 3720  
Db TGAATTTCTTTTCAATAGATTTCTTGGAAAGATGTAGAAAAAATGTGGACACCTATAATC 3720  
QY 3721 CATGGATAGGCTTTTCAGTAGCTGTTTATATCCCTTTACCTGCAAGTGTGTACTGTGATTT 3780  
Db CATGGATAGGCTTTTCAGTAGCTGTTTATATCCCTTTACCTGCAAGTGTGTACTGTGATTT 3780  
QY 3781 TCCCTTTCAATATCATGAGAAAAAATATGAGGAGCAGATCAATTAAGAAAGATCCCTTTT 3840  
Db TCCCTTTCAATATCATGAGAAAAAATATGAGGAGCAGATCAATTAAGAAAGATCCCTTTT 3840  
QY 3841 TCAGAAACCTTTCAACGAAAGTCTTAAATATAGGAAGCTTCCAGAACCCAGACCAATGAG 3900  
Db TCAGAAACCTTTCAACGAAAGTCTTAAATATAGGAAGCTTCCAGAACCCAGACCAATGAG 3900  
QY 3901 ATGAAGATGAAGTCTTCAAGCTGAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 3960  
Db ATGAAGATGAAGTCTTCAAGCTGAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTT 3960  
QY 3961 GTTGTGAGGAGAAACCATCCATTTATGCTCAGCAATTTGTCATAAAGAAATGATGACAA 4020  
Db GTTGTGAGGAGAAACCATCCATTTATGCTCAGCAATTTGTCATAAAGAAATGATGACAA 4020  
QY 4021 AAGATTTTCTTCTTTCAAGAAAAAGTAAAGAAAGTGGCAACTTAAATCACTCTTCTCTG 4080  
Db AAGATTTTCTTCTTTCAAGAAAAAGTAAAGAAAGTGGCAACTTAAATCACTCTTCTCTG 4080  
QY 4081 TGAAGAAAGGAGATCTTAGGACTATTTGGTCCAAATGGTCTGCGCAAAAGCAATTA 4140  
Db TGAAGAAAGGAGATCTTAGGACTATTTGGTCCAAATGGTCTGCGCAAAAGCAATTA 4140  
QY 4141 TTAATATTTCTGTTTGTGTATTTGAAACCACTTCAAGCCAGCTATTTTATGAGGATTT 4200  
Db TTAATATTTCTGTTTGTGTATTTGAAACCACTTCAAGCCAGCTATTTTATGAGGATTT 4200  
QY 4201 CTTCAGAGCAAGTGAAGTATGATTTCACTGAAGTGTATGGTACTGTCTCAGATAA 4260  
Db CTTCAGAGCAAGTGAAGTATGATTTCACTGAAGTGTATGGTACTGTCTCAGATAA 4260  
QY 4261 ACCCTTTTGGGCCAGATCTACTCATTTGCAGGAACATTTTGAATTTTATGAGCTGTCAA 4320  
Db ACCCTTTTGGGCCAGATCTACTCATTTGCAGGAACATTTTGAATTTTATGAGCTGTCAA 4320  
QY 4321 GAATGAGTCAAGTGAAGTGAAGTCAATGATGATGATGATGATGATGATGATGATGATG 4380  
Db GAATGAGTCAAGTGAAGTGAAGTCAATGATGATGATGATGATGATGATGATGATGATG 4380  
QY 4381 AAGAACATCTTCAGAAAGTGTAAAGAAACTACCTGCGAGGAATCAAAAGAAAGTGTGTT 4440  
Db AAGAACATCTTCAGAAAGTGTAAAGAAACTACCTGCGAGGAATCAAAAGAAAGTGTGTT 4440  
QY 4441 TTGCTCTAAGTATGCTAGGGAATCCCTCAGATTTACTTTGCTAGATGAACCATCTACAG 4500  
Db TTGCTCTAAGTATGCTAGGGAATCCCTCAGATTTACTTTGCTAGATGAACCATCTACAG 4500  
QY 4501 TGGATCCCAAGCCAAACAGACATGTGGCAGCAATTCGAACTGCTGATTTTAAAAACAGAA 4560

4501	TGGATCCCAAGAAGCCAAACAGACGACATGTGGCGAGCAATTCGAACTGCAATTTAAAAACAGAA	4561
4561	AGCGGGCTGCTATTTCTGACCACTCACTATATGAGGAGGAGGAGGCTGTCTGTGATCGAG	4620
4561	AGCGGGCTGCTATTTCTGACCACTCACTATATGAGGAGGAGGAGGCTGTCTGTGATCGAG	4620
4621	TAGCTATCATGGTGTCTGGCGAGTTAGATAGTATCGGAACAGTACACACTCTAAAGAGTA	4680
4621	TAGCTATCATGGTGTCTGGCGAGTTAGATAGTATCGGAACAGTACACACTCTAAAGAGTA	4680
4681	AATTTGGAAAAGGCTACTTTTTTGGAAATTAATTTGAAAGGACTGGATAGAAAACTTGAAG	4740
4681	AATTTGGAAAAGGCTACTTTTTTGGAAATTAATTTGAAAGGACTGGATAGAAAACTTGAAG	4740
4741	TAGACCGCTTCAAGAGGAATTCAGTATATTTTCCAAATGCAAGCCGTCAGGAAGTT	4800
4741	TAGACCGCTTCAAGAGGAATTCAGTATATTTTCCAAATGCAAGCCGTCAGGAAGTT	4800
4801	TTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCAACATCTT	4860
4801	TTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCAACATCTT	4860
4861	TTTTTTAAGCTGGAAGAGCTTAACATGCTTTTGGCATTGGAAGATATAGCTTTTCTCAAG	4920
4861	TTTTTTAAGCTGGAAGAGCTTAACATGCTTTTGGCATTGGAAGATATAGCTTTTCTCAAG	4920
4921	CAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGGAAGATAATAGTT	4980
4921	CAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGGAAGATAATAGTT	4980
4981	GTGGAACTTTAAACAGACACACTTTGTGGGAGCAACACAGAGAGATAGAGTAGTATTTT	5040
4981	GTGGAACTTTAAACAGACACACTTTGTGGGAGCAACACAGAGAGATAGAGTAGTATTTT	5040
5041	GAATTTGTATGTTTCGGTCTGCTTACTGCGACTTTCTTTTCTTTTCACTTAAATTTTAACTT	5100
5041	GAATTTGTATGTTTCGGTCTGCTTACTGCGACTTTCTTTTCTTTTCACTTAAATTTTAACTT	5100
5101	TGGTTTTAAAAAGTTTTTTATTTGGAATGGTAACTGGAGAACCAAGAACGCACCTTGAATTTT	5160
5101	TGGTTTTAAAAAGTTTTTTATTTGGAATGGTAACTGGAGAACCAAGAACGCACCTTGAATTTT	5160
5161	TTCTAAGCTCCTTAATTTGAAATGCTGGTGTGTGTTTCTTTTCTTTTCTTTTAAATAAAAACG	5220
5161	TTCTAAGCTCCTTAATTTGAAATGCTGGTGTGTGTTTCTTTTCTTTTCTTTTAAATAAAAACG	5220
5221	TATGTTAAATTAAGTGAAGCTGCATGTTTGTATTTGAAGTATATTGGAACATATAGATTGTT	5280
5221	TATGTTAAATTAAGTGAAGCTGCATGTTTGTATTTGAAGTATATTGGAACATATAGATTGTT	5280
5281	ATGTCATCTTTTCCACCAATTCAGAAAAAGTGCCTTCTGAATTTGTGATTTAAAGGAATTTG	5340
5281	ATGTCATCTTTTCCACCAATTCAGAAAAAGTGCCTTCTGAATTTGTGATTTAAAGGAATTTG	5340
5341	AATAGAAATAGTTTTATTTTAAAGTTATCTTTTAAGTTTATGCGATCTTCTTTAAATTAAGTAC	5400
5341	AATAGAAATAGTTTTATTTTAAAGTTATCTTTTAAGTTTATGCGATCTTCTTTAAATTAAGTAC	5400
5401	GTAATGTTCCCAATCTAAATAAAAAAACAATTAATTCATAACTAAAGCATAGAAAAAGATACATAA	5460
5401	GTAATGTTCCCAATCTAAATAAAAAAACAATTAATTCATAACTAAAGCATAGAAAAAGATACATAA	5460
5461	AGCAATGTGAAAAGTT	5475
5461	AGCAATGTGAAAAGTT	5475

RESIST 2

RESULT 2  
US-10-005-338B-1

US-10-005-338B-1 : Sequence 1. Application US/10005338B

; Sequence 1, Application US/1000  
; Publication No. US20030044895A1

; Publication No. US20  
: GENERAL INFORMATION:

; GENERAL INFORMATION:  
: APPLICANT: DENEFFLE, Patrice



Db 1594 AGGAGCTGGAGTCAACTAAAGCTGTTATATGGGAGAAACCTGCTGTTGTAGAAATAGATA 1653  
Qy 757 CCTTCCCGGAGGAGTAATTTTATATATACCTAGTATAGCATTTTACCTTTTGGTACT 816  
Db 1654 CCTTCCCGGAGGAGTAAATTTTATATATACCTAGTATAGCATTTTACCTTTTGGTACT 1713  
Qy 817 TTTTGGCAATTCATATCGTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAA 876  
Db 1714 TTTTGGCAATTCATATCGTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAA 1773  
Qy 877 TGGGACATTCATGATACCTGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTT 936  
Db 1774 TGGGACATTCATGATACCTGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTTGGCTTTT 1833  
Qy 937 TTTTATGTCCTTTCTTTATGGCAGTCAATGGCAGAGCTTCTTTGTTATTTTCTTCAAAGTA 996  
Db 1834 TTTTATGTCCTTTCTTTATGGCAGTCAATGGCAGAGCTTCTTTGTTATTTTCTCAAAGTA 1893  
Qy 997 GCAGCATTTGATATTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTT 1056  
Db 1894 GCAGCATTTGATATTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTT 1953  
Qy 1057 TAATGCTGACACCTCTTTTAAATAAATCAAAACATGTGGGAATAGTTGAAATTTTGTGA 1116  
Db 1954 TAATGCTGACACCTCTTTTAAATAAATCAAAACATGTGGGAATAGTTGAAATTTTGTGA 2013  
Qy 1117 CTGTGGCTTTTGGATTTATTTGGCTTTATGATTAATTCCTCATAGAAAAGTTTCCCAATCGT 1176  
Db 2014 CTGTGGCTTTTGGATTTATTTGGCTTTATGATTAATTCCTCATAGAAAAGTTTCCCAATCGT 2073  
Qy 1177 TAGTGTGGCTTTTCAAGTCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTT 1236  
Db 2074 TAGTGTGGCTTTTCAAGTCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTTCTGCTTTT 2133  
Qy 1237 TGCATTTAGAGATTTTAAATGAAGTGTCTTCAATTTTCAAAATTTGACATGCGGCCCCATC 1296  
Db 2134 TGCATTTAGAGATTTTAAATGAAGTGTCTTCAATTTTCAAAATTTGACATGCGGCCCCATC 2193  
Qy 1297 CTCTAATTTATACAAATATCATGCTCACTTAATAGTATATTTCTATGCTCTTGGCTG 1356  
Db 2194 CTCTAATTTATACAAATATCATGCTCACTTAATAGTATATTTCTATGCTCTTGGCTG 2253  
Qy 1357 TCTATCTTGATCAAGTCAATCCAGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTC 1416  
Db 2254 TCTATCTTGATCAAGTCAATCCAGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTC 2313  
Qy 1417 TGAAGCTTTCATATTTGGTCAAAAGAGCAAAAGAAATTAAGGAGTATACAGAGGCAATG 1476  
Db 2314 TGAAGCTTTCATATTTGGTCAAAAGAGTAAAGAAATTAAGGAGTATACAGAGGCAATG 2373  
Qy 1477 TTAATGGAAATATTAGTTTATGTAATTTAGTCAAGCTTTCTTCAAGAAATTTGTAGGAA 1536  
Db 2374 TTAATGGAAATATTAGTTTATGTAATTTAGTCAAGCTTTCTTCAAGAAATTTGTAGGAA 2433  
Qy 1537 AAGAAGCCATTAAGAAATTTAGTGGTATTCAGAGACATACAGAAAGAGGGTGAATAATGTGG 1596  
Db 2434 AAGAAGCCATTAAGAAATTTAGTGGTATTCAGAGACATACAGAAAGAGGGTGAATAATGTGG 2493  
Qy 1597 AGGCTTTGAGAAATTTGCTATTCACATATATGAGGCTCAGATTTACCTTACTTTGGCC 1656  
Db 2494 AGGCTTTGAGAAATTTGCTATTCACATATATGAGGCTCAGATTTACCTTACTTTGGCC 2553  
Qy 1657 ACAGTGGAAAGGAAAGATGATTTGATGAATTTCTTTGTTGACTCTGCCCCCTTTCTG 1716  
Db 2554 ACAGTGGAAAGGAAAGATGATTTGATGAATTTCTTTGTTGACTCTGCCCCCTTTCTG 2613  
Qy 1717 ATGGGTTTGCATCTATATATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 1776  
Db 2614 ATGGGTTTGCATCTATATATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 2673  
Qy 1777 GAAAAATGATTTGGCATTTGTCCAGATTTAGATATACACTTTGATTTTGCACAGTAGAAG 1836  
Db 2674 GAAAAATGATTTGGCATTTGTCCAGATTTAGATATACACTTTGATTTTGCACAGTAGAAG 2733

Qy 1837 ABAATTTTCAATTTTGGCTTCAATCAAGGATACACAGCCAAACAATATATACAAAG 1896  
Db 2734 ABAATTTTCAATTTTGGCTTCAATCAAGGATACACAGCCAAACAATATATACAAAG 2793  
Qy 1897 TGCAGAGGTTTACTAGATTTAGACATTCGACATPATCAAGATAAACCAAGCTAAAAAT 1956  
Db 2794 TGCAGAGGTTTACTAGATTTAGACATTCGACATPATCAAGATAAACCAAGCTAAAAAT 2853  
Qy 1957 TAAAGTGGTGCACAAAAGAAAGCTGCTATAGGAATTTGCTTTCTGGGAACCCAAAGA 2016  
Db 2854 TAAAGTGGTGCACAAAAGAAAGCTGCTATAGGAATTTGCTTTCTGGGAACCCAAAGA 2913  
Qy 2017 TACTGCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGA 2076  
Db 2914 TACTGCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGA 2973  
Qy 2077 ATCTTTTAAATACAGAAAAGCCATCGGCTGACAGTGTTCAGTACTCATTTTCATGGATG 2136  
Db 2974 ATCTTTTAAATACAGAAAAGCCAAATCGGCTGACAGTGTTCAGTACTCATTTTCATGGATG 3033  
Qy 2137 AAGCTGACATTTCTTCAGATAGGAAAGCTGTGATATCAAGGAATGCTGAAATGTTG 2196  
Db 3034 AAGCTGACATTTCTTCAGATAGGAAAGCTGTGATATCAAGGAATGCTGAAATGTTG 3093  
Qy 2197 GTTCTTCAATGTTCTTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
Db 3094 GTTCTTCAATGTTCTTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153  
Qy 2257 ACAAAATTTTGCCACAGAACTCTTTTCTCACTGGTTAAACAAATATACCTGAGGCTA 2316  
Db 3154 ACAAAATTTTGCCACAGAACTCTTTTCTCACTGGTTAAACAAATATACCTGAGGCTA 3213  
Qy 2317 CTTTATTAACAGAAATGACCAACAACTTGTGTATAGCTTGTGCTTTCAAGGACATGGACA 2376  
Db 3214 CTTTATTAACAGAAATGACCAACAACTTGTGTATAGCTTGTGCTTTCAAGGACATGGACA 3273  
Qy 2377 AATTTTCAAGTGTGTTTCTGCTTACAGTCAATTTCAAAATTTGGGTGCTATTTCTTATG 2436  
Db 3274 AATTTTCAAGTGTGTTTCTGCTTACAGTCAATTTCAAAATTTGGGTGCTATTTCTTATG 3333  
Qy 2437 GTGTTTCCATGACGACTTTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATG 2496  
Db 3334 GTGTTTCCATGACGACTTTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATG 3393  
Qy 2497 ACCAGCAGATTTATAGTGTATTTTACTCAGCAGCCACTCGAGGAGAAATGGATTCAAAAT 2556  
Db 3394 ACCAGCAGATTTATAGTGTATTTTACTCAGCAGCCACTCGAGGAGAAATGGATTCAAAAT 3453  
Qy 2557 CTTTGTGATGAATGGAACAGAGCTTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
Db 3454 CTTTGTGATGAATGGAACAGAGCTTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 3513  
Qy 2617 GCACATGAGCCTTTGGAAAACACAGATGTATACATAGCAAGTTTCATTTCTTTTACT 2676  
Db 3514 GCACATGAGCCTTTGGAAAACACAGATGTATACATAGCAAGTTTCATTTCTTTTACT 3573  
Qy 2677 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736  
Db 3574 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 3633  
Qy 2737 TTCAGATTTTATGTTTGGTTTCATCATCTTTTAAATAATGCTGTGTTCCCATCAAC 2796  
Db 3634 TTCAGATTTTATGTTTGGTTTCATCATCTTTTAAATAATGCTGTGTTCCCATCAAC 3693  
Qy 2797 TTGTTTCCAGCTTATATTTTCTAAAACTGGAGACAAACCAATATAAACAAGTCAAGTGC 2856  
Db 3694 TTGTTTCCAGCTTATATTTTCTAAAACTGGAGACAAACCAATATAAACAAGTCAAGTGC 3753  
Qy 2857 TGCCTTCTTCAAAATTTCTGCTGACTCAGATATCAGTGTCTTTATAGCTTTTTCACAGCC 2916  
Db 3754 TGCCTTCTTCAAAATTTCTGCTGACTCAGATATCAGTGTCTTTATAGCTTTTTCACAGCC 3813

QY 2917 AGAACAATATGTCAGATGATTAATGACAGTACTATGATCCGTGCTCCCATAGTG 2976  
 Db 3814 AGAACAATATGTCAGATGATTAATGACAGTACTATGATCCGTGCTCCCATAGTG 3873  
 QY 2977 CGGCTTTAAATGTCAGATGATTAATGACAGTACTATGATCCGTGCTCCCATAGTG 3036  
 Db 3874 CGGCTTTAAATGTCAGATGATTAATGACAGTACTATGATCCGTGCTCCCATAGTG 3933  
 QY 3037 CTATGCTTTATCTTTTACCTATTAATGAGTAAATGATTAATGAGTAAATGATTAAT 3096  
 Db 3934 CTATGCTTTATCTTTTACCTATTAATGAGTAAATGATTAATGAGTAAATGATTAAT 3993  
 QY 3097 TAAATGTCAGTGAACCATCCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 3156  
 Db 3994 TAAATGTCAGTGAACCATCCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 4053  
 QY 3157 TAGTTTTTAAATGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3216  
 Db 4054 TAGTTTTTAAATGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4113  
 QY 3217 TGGCAGCTTACCTTGGCATGGAATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 3276  
 Db 4114 TGGCAGCTTACCTTGGCATGGAATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 4173  
 QY 3277 TTAATGTCAGTGAACCATCCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 3336  
 Db 4174 TTAATGTCAGTGAACCATCCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 4233  
 QY 3337 CCTTATTTTATCACTTATTTGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 3396  
 Db 4234 CCTTATTTTATCACTTATTTGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 4293  
 QY 3397 TATATTTTATCACTTATTTGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 3456  
 Db 4294 TATATTTTATCACTTATTTGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 4353  
 QY 3457 CAGTTATTTGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3516  
 Db 4354 CAGTTATTTGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4413  
 QY 3517 AATTTTGGTCACTTATTTGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 3576  
 Db 4414 AATTTTGGTCACTTATTTGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 4473  
 QY 3577 CTTTCTTTATGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3636  
 Db 4474 CTTTCTTTATGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4533  
 QY 3637 CAATCTATCCATTTGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3696  
 Db 4534 CAATCTATCCATTTGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4593  
 QY 3697 GAAAAAATGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3756  
 Db 4594 GAAAAAATGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4653  
 QY 3757 ACCTGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3816  
 Db 4654 ACCTGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4713  
 QY 3817 GATCAATAAGAAAGATCCCTTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3876  
 Db 4714 GATCAATAAGAAAGATCCCTTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4773  
 QY 3877 TTCAGAAACCCAGCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3936  
 Db 4774 TTCAGAAACCCAGCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4833  
 QY 3937 TCAAGAGCTGATGGGTGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 3996  
 Db 4834 TCAAGAGCTGATGGGTGTCAGTGAATTTTCAAGCAGCTTTGCTTGAATCATTTGATGCA 4893  
 QY 3997 TGCATAAAGATATGATGACAGAAAGATTTTCTTCTTTCAAGAAAGTAAAGAGTGG 4056

Db 4894 TGCATAAAGATATGATGACAGAAAGATTTTCTTCTTTCAAGAAAGTAAAGAGTGG 4953  
 QY 4057 CAACTAAATACATCTCTTTCTGTCGAAAAAGGAGAGATCTTAGGACTATTTGGGTCCAA 4116  
 Db 4954 CAACTAAATACATCTCTTTCTGTCGAAAAAGGAGAGATCTTAGGACTATTTGGGTCCAA 5013  
 QY 4117 ATGGTCTGGCAAAAGACAAATTTAATTAATTTCTGTCGTCGTCGTCGTCGTCGTCGTCG 4176  
 Db 5014 ATGGTCTGGCAAAAGACAAATTTAATTAATTTCTGTCGTCGTCGTCGTCGTCGTCGTCG 5073  
 QY 4177 GCCAGGTATTTTATGAGAGATTTCTTTCAGAGACAGTGAAGATGATGATTTCACTGAAGT 4236  
 Db 5074 GCCAGGTATTTTATGAGAGATTTCTTTCAGAGACAGTGAAGATGATGATTTCACTGAAGT 5133  
 QY 4237 GTATGGTATCTGTCCTCAGATAAACCTTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4296  
 Db 5134 GTATGGTATCTGTCCTCAGATAAACCTTTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5193  
 QY 4297 TTCAAAATTTATGAGCTGTCAAAAGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4356  
 Db 5194 TTCAAAATTTATGAGCTGTCAAAAGGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5253  
 QY 4357 GAATAACACATGACCTTTGATTTTAAAGAACATCTTTCAGAGACCTGTCGTCGTCGTCGTCG 4416  
 Db 5254 GAATAACACATGACCTTTGATTTTAAAGAACATCTTTCAGAGACCTGTCGTCGTCGTCGTCG 5313  
 QY 4417 CAGGATCAAAACGAAAGTGTGTCCTTAAGTATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4476  
 Db 5314 CAGGATCAAAACGAAAGTGTGTCCTTAAGTATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5373  
 QY 4477 TGCATAGTAACCATCTACAGGTATGATCCCAAGACCAAGACCAAGACCAAGACCAAGACCA 4536  
 Db 5374 TGCATAGTAACCATCTACAGGTATGATCCCAAGACCAAGACCAAGACCAAGACCAAGACCA 5433  
 QY 4537 TTGCAATGTCATTTTAAAGAACGAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4596  
 Db 5434 TTGCAATGTCATTTTAAAGAACGAGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5493  
 QY 4597 AGGCAGAGGCTGTCTGTGATCGAGTATGATCGAGTATGATCGAGTATGATCGAGTATGATCG 4656  
 Db 5494 AGGCAGAGGCTGTCTGTGATCGAGTATGATCGAGTATGATCGAGTATGATCGAGTATGATCG 5553  
 QY 4657 GACAGTACCAATCTTAAGAGTAAATTTGGAAGGCTACTTTTGGAAATTAATTTGA 4716  
 Db 5554 GACAGTACCAATCTTAAGAGTAAATTTGGAAGGCTACTTTTGGAAATTAATTTGA 5613  
 QY 4717 AGGACTGATAGAAACCTTAGAAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTTCC 4776  
 Db 5614 AGGACTGATAGAAACCTTAGAAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTTCC 5673  
 QY 4777 CAAATGCAAGCCGTGAGAAAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4836  
 Db 5674 CAAATGCAAGCCGTGAGAAAGTGTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5733  
 QY 4837 ATGTTTCACTCCCTTCAATCTTTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4896  
 Db 5734 ATGTTTCACTCCCTTCAATCTTTTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 5793  
 QY 4897 TTGAAGAATATAGCTTTTCTCAAGAACATTTGGAACAGGTTTGTAGAACTCACTAAAG 4956  
 Db 5794 TTGAAGAATATAGCTTTTCTCAAGAACATTTGGAACAGGTTTGTAGAACTCACTAAAG 5853  
 QY 4957 AACAGAGGAGGAGATTAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5016  
 Db 5854 AACAGAGGAGGAGATTAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5913  
 QY 5017 CACAAGAGATGAGTATGATTTTGAATTTGTATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5076  
 Db 5914 CACAAGAGATGAGTATGATTTTGAATTTGTATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 5973  
 QY 5077 TTTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGGTTTTTATTTGGAATGGAATGGAATGGA 5136

D <sub>b</sub>	5974	TTCTTTTTCACCTTAATTTTAACTCTTTGGTTTAAAAAGTTTTTTTATTGGATCGTAACCTGGA	6033
Q <sub>y</sub>	5137	GAACCAAGAACGCACCTTGAAATTTTCTTAAGCTCTTAATTAAGAAATGCTGTGGTTGTGTG	5196
D <sub>b</sub>	6034	GAACCAAGAACGCACCTTGAAATTTTCTTAAGCTCTTAATTAAGAAATGCTGTGGTTGTGTG	6093
Q <sub>y</sub>	5197	TTTTGCTTTTCTTTTAAATATAAACAAGTATGATATAAATTAAGTGAAGCTGCATGTTTCTGATTGA	5256
D <sub>b</sub>	6094	TTTTGCTTTTCTTTTAAATATAAACAAGTATGATATAAATTAAGTGAAGCTGCATGTTTCTGATTGA	6153
Q <sub>y</sub>	5257	AGTATATTGAACATAATAGTTTGTATGTCATCTTTTTCACCAATTCAGAAAAAGTCTTCT	5316
D <sub>b</sub>	6154	AGTATATTGAACATAATAGTTTGTATGTCATCTTTTTCACCAATTCAGAAAAAGTCTTCT	6213
Q <sub>y</sub>	5317	GAATTTGTGATTTTAAAGGAATTTGAAATAGATAGTATTTTATTATTTTAAGTTATCTTTTAAGTT	5376
D <sub>b</sub>	6214	GAATTTGTGATTTTAAAGGAATTTGAAATAGATAGTATTTTATTATTTTAAGTTATCTTTTAAGTT	6273
Q <sub>y</sub>	5377	TATGCCATCTTCTTTTAAATAAGTACGTAAATGTTCCAATCTAAATAAAAAAACTAAAT	5430
D <sub>b</sub>	6274	TATGCCATCTTCTTTTAAATAAGTACGTAAATGTTCCAATCTAAATAAAAAAACTAAAT	6327

### RESULT 3

```

US-09-971-121-5
; Sequence 5, Application US/09971121
; Patent No. US2002011477A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Nepomnichy, Boris
; TITLE OF INVENTION: No. US2002011477A1el Human Transporter Proteins and Polynucleot
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0250-USA
; CURRENT APPLICATION NUMBER: US/09/971.121
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/239,629
; PRIOR FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 5262
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-971-121-5

```

Query Match	85.2%;	Score 4664;	DB 9;	Length 5262;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 5014;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
Qy	97	AGGTTTATTCAGAAAACATGTCCACTGCAATTTAGGAGGTAGGAGTTTGGAGACAGACCA	156	
Db	238	AGGTTTATTCAGAAAACATGTCCACTGCAATTTAGGAGGTAGGAGTTTGGAGACAGACCA	297	
Qy	157	GAAACACTTCTACTGAAGAATTTACTTTAATTAATGCAGAACCAAAAAGTAGTGTTCAGG	216	
Db	298	GAAACACTTCTACTGNAGAAATTTACTTTAATTAATGCAGAACCAAAAAGTAGTGTTCAGG	357	
Qy	217	AAATTCCTTTTCCACTATTTTTTTATTTTGGTTAAATTAATTTAGCAATGATGCATCAA	276	
Db	358	AAATTCCTTTTCCACTATTTTTTTATTTTGGTTAAATTAATTTAGCAATGATGCATCAA	417	
Qy	277	ATAAGAAATATGAGAAAGTGCTTAATATAGAACTCAATCCCTATGGACAAAGTTTACTCTTT	336	
Db	418	ATAAGAAATATGAGAAAGTGCTTAATATAGAACTCAATCCCTATGGACAAAGTTTACTCTTT	477	
Qy	337	CTAATCTCAATTCCTTGGATATATCTCCAGTGCATTAATATTACAAGCAGCATCATGCAGAAAG	396	
Db	478	CTAATCTCAATTCCTTGGATATATCTCCAGTGCATTAATATTACAAGCAGCATCATGCAGAAAG	537	
Qy	337	TGTCCTACTGATCATCTA CCTGATGTCATAATTTA CTGAAAGAAATATACAAATGAAAAGAAA	456	
Db	538	TGTCCTACTGATCATCTA CCTGATGTCATAATTTA CTGAAAGAAATATACAAATGAAAAGAAA	597	

1678 AAGAAAGCCATAAGAAATTAAGTGTATTCAGAGACACATACAGAAAGAGGTGAAATGTGG 1737  
1597 AGGCTTTGAGAAATTTGTCATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCC 1656  
1738 AGGCTTTGAGAAATTTGTCATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCC 1797  
1657 ACAGTGGAAACAGAAAGAGTACATTGATGAATATTTCTTTGGGACTCTGCCCACCTTCTG 1716  
1798 ACAGTGGAAACAGAAAGAGTACATTGATGAATATTTCTTTGGGACTCTGCCCACCTTCTG 1857  
1717 ATGGGTTTGGCATATATATGACACAGAGTCTCAGAAATAGATCAATGTTTGAAGCAA 1776  
1858 ATGGGTTTGGCATATATATGACACAGAGTCTCAGAAATAGATCAATGTTTGAAGCAA 1917  
1777 GAAATATGATGGCAATTTGTCCACAGTTAGATATACACTTTTGATGTTTGGACAGTAGAAG 1836  
1918 GAAATATGATGGCAATTTGTCCACAGTTAGATATACACTTTTGATGTTTGGACAGTAGAAG 1977  
1837 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATATATACAGAAAG 1896  
1978 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATATATACAGAAAG 2037  
1897 TGCAGAGGTTTACTAGATTAGACATGCGAGACTATCAAGAGATAACCAAGCTTAAAAAT 1956  
2038 TGCAGAGGTTTACTAGATTAGACATGCGAGACTATCAAGAGATAACCAAGCTTAAAAAT 2097  
1957 TAAGTGGTGGTCAAAAAGAAAGCTGTCAATTAGGAATGCTGTTCTGGGAACCCAAAGA 2016  
2098 TAAGTGGTGGTCAAAAAGAAAGCTGTCAATTAGGAATGCTGTTCTGGGAACCCAAAGA 2157  
2017 TACTGCTGTAGATGAAACCAAGCTGGAATGGACCCCTGTCTCGACATATTTATATGGA 2076  
2158 TACTGCTGTAGATGAAACCAAGCTGGAATGGACCCCTGTCTCGACATATTTATGGA 2217  
2077 ATCTTTTAAATPACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATG 2136  
2218 ATCTTTTAAATPACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATG 2277  
2137 AAGCTGACATTTCCAGATAGGAAGCTGTGATATCAAGAGGAATGCTGAAATGTGTG 2196  
2278 AAGCTGACATTTCCAGATAGGAAGCTGTGATATCAAGAGGAATGCTGAAATGTGTG 2337  
2197 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
2338 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2397  
2257 ACAATATTTGCGCACAGATCTCTTCTTCACTGGTTAAACACATATACCTGGAGCTA 2316  
2398 ACAATATTTGCGCACAGATCTCTTCTTCACTGGTTAAACACATATACCTGGAGCTA 2457  
2317 CTTTATTACACAGAAATGACCAACAACTTGTGTATAGCTTGCCTTTCAAGGACATGGACA 2376  
2458 CTTTATTACACAGAAATGACCAACAACTTGTGTATAGCTTGCCTTTCAAGGACATGGACA 2517  
2377 AATTTTCAGGTTTGTGTCGCCCTAGACAGTCAATCAAAATGGGGTGTCAATTTCTTATG 2436  
2518 AATTTTCAGGTTTGTGTCGCCCTAGACAGTCAATCAAAATGGGGTGTCAATTTCTTATG 2577  
2437 GTGTTTCCATGACGACTTTGGAAGACGTATTTTAAAGCTAGAAAGTTGAAGCAAAATG 2496  
2578 GGGTTTCCATGACGACTTTGGAAGACGTATTTTAAAGCTAGAAAGTTGAAGCAAAATG 2637  
2497 ACCAAGCAGATPATAGTGTATTTACTCAGCAGCCACTGGAGGAAGAAATGANTTCAAAAT 2556  
2638 ACCAAGCAGATPATAGTGTATTTACTCAGCAGCCACTGGAGGAAGAAATGANTTCAAAAT 2697  
2557 CTTTGTGAAATGGAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
2698 CTTTGTGAAATGGAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 2757  
2617 GCACCATGAGCCTTTGGAAACCAACAGATGTATACAAATAGCAAGTTTCATTTCTTACCT 2876

2758 GCACCATGAGCCTTTGGAAACCAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTACCT 2817  
2677 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736  
2818 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2877  
2737 TGCAGATTTTAAATGTTTGGTTTGCATCACTCTTTTAAAAATGCTGTGGTCCCAACAAC 2796  
2878 TGCAGATTTTAAATGTTTGGTTTGCATCACTCTTTTAAAAATGCTGTGGTCCCAACAAC 2937  
2797 TGTGTTCCAGACTTATATTTTCTTAAACCTGGAGACAAACCAATATAACAAACAGTTC 2856  
2938 TGTGTTCCAGACTTATATTTTCTTAAACCTGGAGACAAACCAATATAACAAACAGTTC 2997  
2857 TGTGTTCTTCAAAATTCGCTGACTCAGATATCAGTATCAGTATCTTATTAGCTTTTTCACAAGCC 2916  
2998 TGTGTTCTTCAAAATTCGCTGACTCAGATATCAGTATCAGTATCTTATTAGCTTTTTCACAAGCC 3057  
2917 AGAACATATATGTCGAGATGATTAAATGACAGTATCTATGATCCGTGGCTCCCATATAGT 2976  
3058 AGAACATATATGTCGAGATGATTAAATGACAGTATCTATGATCCGTGGCTCCCATATAGT 3117  
2977 CGGCTTTAAATGTGATGCAATTCAGAAAGGACTATGTTTTTCGAGCTGTTTTCAACAGTA 3036  
3118 CGGCTTTAAATGTGATGCAATTCAGAAAGGACTATGTTTTTCGAGCTGTTTTCAACAGTA 3177  
3037 CTATGTTTATTTTCTTTTACCTATATTTAGTGAATATCAATTAGTAACTACTATCTTTATCAT 3096  
3178 CTATGTTTATTTTCTTTTACCTATATTTAGTGAATATCAATTAGTAACTACTATCTTTATCAT 3237  
3097 TAAATGTGATGCAAAACCATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTTACTGATA 3156  
3238 TAAATGTGATGCAAAACCATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTTACTGATA 3297  
3157 TAGTTTTTAAATTTGAGCTGTATTTTCAAGAGCTTTGCTTGGATCATTTGTTACTGCA 3216  
3298 TAGTTTTTAAATTTGAGCTGTATTTTCAAGAGCTTTGCTTGGATCATTTGTTACTGCA 3357  
3217 TGCACCTTTACTTTTGCATGGAATTCAGAGAAATCAAAAGATCAAAAGCTTATATCTCAAC 3276  
3358 TGCACCTTTACTTTTGCATGGAATTCAGAGAAATCAAAAGATCAAAAGCTTATVACTCAAC 3417  
3277 TTAACCTTTCAGTCTTTTGCATCTGATATTTGATTTGGATGGACAGCTGTTTGTATATCC 3336  
3418 TTAACCTTTCAGTCTTTTGCATCTGATATTTGATTTGGATGGACAGCTGTTTGTATATCC 3477  
3337 CCTTATTTTATCATTTCTTATTTGATGCTAGGAAGCTTATTTGGCATTTCAATTATGGAT 3396  
3478 CCTTATTTTATCATTTCTTATTTGATGCTAGGAAGCTTATTTGGCATTTCAATTATGGAT 3537  
3397 TATATTTTATATCTGTAAGTTTCTTGTGCTGCTGTTTTTGGCTTATTTGGTATGTTCCAT 3456  
3538 TATATTTTATATCTGTAAGTTTCTTGTGCTGCTGTTTTTGGCTTATTTGGTATGTTCCAT 3597  
3457 CAGTATTTCTGTTCTCATTTATTTGCTTCTTCAACCTTTAAGAAATTTTAAATATCAAAAG 3516  
3598 CAGTATTTCTGTTCTCATTTATTTGCTTCTTCAACCTTTAAGAAATTTTAAATATCAAAAG 3657  
3517 AATTTTGGTCAATTTATCTTATTTCTGTCGAGCTGTTGGCTTGTATTTGCAATCACTGAAATAA 3576  
3658 AATTTTGGTCAATTTATCTTATTTCTGTCGAGCTGTTGGCTTGTATTTGCAATCACTGAAATAA 3717  
3577 CTTTCTTTATGGATACACAAATTTGCAATTTCTTCAATTTGCTTGTATTCATCATCTTC 3636  
3718 CTTTCTTTATGGATACACAAATTTGCAATTTCTTCAATTTGCTTGTATTCATCATCTTC 3777  
3637 CAATCTATCCACTTCTAGGTTGCTGATTTCTTTTCAATAGAAATTTCTTGGAAAGATGTAC 3696  
3778 CAATCTATCCACTTCTAGGTTGCTGATTTCTTTTCAATAGAAATTTCTTGGAAAGATGTAC 3837  
3697 GAAAAAATGTGACACCTATATTCATGGATAGGCTTTTCACTAGTGTATATGCTT 3756  
3838 GAAAAAATGTGACACCTATATTCATGGATAGGCTTTTCACTAGTGTATATGCTT 3897

Qy	3757	ACCTGCAGTGTACTGTGGATTTTTCTCTTACAATACTATGAGAAAAATATGGAGCCA	3816
Db	3998	ACCTGCAGTGTGTACTGTGGATTTTCTCTTACAATACTATGAGAAAAATATGGAGCCA	3957
Qy	3817	GATCAATAAGAAAAAGATGCCCTTTTTCAGAAACCTTTTCAAAGAACTTTAAAAATAGGAAGC	3876
Db	3958	GATCAATAAGAAAAAGATGCCCTTTTTCAGAAACCTTTTCAAAGAACTTTAAAAATAGGAAGC	4017
Qy	3877	TTCCAGAACCCACAGACAAATGAGATGAGATGAGATGAGATGCAAGAGCTGAAAGACTAAAGG	3936
Db	4018	TTCCAGAACCCACAGACAAATGAGATGAGATGAGATGAGATGCAAGAGCTGAAAGACTAAAGG	4077
Qy	3937	TCAAAGAGCTGATGGGTTGCGAGTGTGTGAGGAGAAACCAATCCATTTATGGTCAGCAATT	3996
Db	4078	TCAAAGAGCTGATGGGTTGCGAGTGTGTGAGGAGAAACCAATCCATTTATGGTCAGCAATT	4137
Qy	3997	TGCATAAAGAAATATGATGA CAAGAAAGATTTTCTTCTTTCAGAAAGATGAAAGAGTGG	4056
Db	4138	TGCATAAAGAAATATGATGACAAAGAAAGATTTTCTTCTTTCAGAAAGAAAGTAAAGAAAGTGG	4197
Qy	4057	CAACTAAATACATCTCTTCTGTGTGAAAAAGGAGAGACTTTAGGACTATTTGGGTCCAA	4116
Db	4198	CAACTAAATACATCTCTTCTGTGTGAAAAAGGAGAGACTTTAGGACTATTTGGGTCCAA	4257
Qy	4117	ATGTGCTGGCAAAAGCACAAATTAATTAATTTCTGGTTGGTGATTTGAAACCAACTTCAG	4176
Db	4258	ATGTGCTGGCAAAAGCACAAATTAATTAATTTCTGGTTGGTGATTTGAAACCAACTTCAG	4317
Qy	4177	GCCAGGTATTTTATAGGAGATTTCTTTCAGAGACAAGTGAAGATGATGATCTACTGAAGT	4236
Db	4318	GCCAGGTATTTTATAGGAGATTTCTTTCAGAGACAAGTGAAGATGATGATCTACTGAAGT	4377
Qy	4237	GTATGGGTTACTGTCTCAGATAAACCCCTTTGTGGCCAGATACTTACATTTCCAGGAAACATT	4296
Db	4378	GTATGGGTTACTGTCTCAGATAAACCCCTTTGTGGCCAGATACTTACATTTCCAGGAAACATT	4437
Qy	4297	TTGAAATTTATGGAGCTGTCAAAGGATGAGTGCAGTGCATGCAAGAGAGTCAATAAGTC	4356
Db	4438	TTGAAATTTATGGAGCTGTCAAAGGATGAGTGCAGTGCATGCAAGAGAGTCAATAAGTC	4497
Qy	4357	GAATAACACATGCATTTGATTTAAAGAAACATCTTTCAGAGAAGCTGTAAGAAAGAACTACCTG	4416
Db	4498	GAATAACACATGCATTTGATTTAAAGAAACATCTTTCAGAGAAGCTGTAAGAAAGAACTACCTG	4557
Qy	4417	CAGAAATCAAACGAAAGTTGTGTTTCTCTTAAGTATGCTTAGGAAATCCTCAGATTACTTT	4476
Db	4558	CAGAAATCAAACGAAAGTTGTGTTTCTCTTAAGTATGCTTAGGAAATCCTCAGATTACTTT	4617
Qy	4477	TGCTAGATGAACCATCTACAGGTATGATCCCAAAGCCAAACAGCACATGTGCGGAGCAA	4536
Db	4618	TGCTAGATGAACCATCTACAGGTATGATCCCAAAGCCAAACAGCACATGTGCGGAGCAA	4677
Qy	4537	TTCGAACTGCATTTAAAAAAGAGAGCGGCTGCTATCTTGACCACTCAGTATGAGAGG	4596
Db	4678	TTCGAACTGCATTTAAAAAAGAGAGCGGCTGCTATCTTGACCACTCAGTATGAGAGG	4737
Qy	4597	AGGCAGAGGCTGCTGTGATCGAGTACTCATGTGTGCTCGGCGAGTTAAGATGTATCG	4656
Db	4738	AGGCAGAGGCTGCTGTGATCGAGTACTCATGTGTGCTCGGCGAGTTAAGATGTATCG	4797
Qy	4657	GACAGGTACCAATCTTAAGAGTAAATTTGGAAAAAGCTACTTTTTGGAAATTTAAATTGA	4716
Db	4798	GACAGGTACCAATCTTAAGAGTAAATTTGGAAAAAGCTACTTTTTGGAAATTTAAATTGA	4857
Qy	4717	AGGACTGATAGAAAAACCTAGAGTAGACCGCCTTCAAAGAGAAATTCAGTATATTTTCC	4776
Db	4858	AGGACTGATAGAAAAACCTAGAGTAGACCGCCTTCAAAGAGAAATTCAGTATATTTTCC	4917
Qy	4777	CAATGCAAGCCGTCAGGAAAGTTTTTCTCTATTTTGGCTTATAAATTCCTAAGGAAG	4836
Db	4918	CAATGCAAGCCGTCAGGAAAGTTTTTCTCTATTTTGGCTTATAAATTCCTAAGGAAG	4977

QY	4837	ATGTTTCAGTCCCTTTTCACAAATCTTTTTTAAAGCTGGAGAAAGCTAAACATGCTTTTGCCA	4898
DB	4978	ATGTTTCAGTCCCTTTTCACAAATCTTTTTTAAAGCTGGAGAAAGCTAAACATGCTTTTGCCA	5037
QY	4897	TTGAAGAAATATAGCTTTTCTCAAGCAACATTGGAAACAGGTTTTTGTAGAACTCACATAAG	4956
DB	5038	TTGAAGAAATATAGCTTTTCTCAAGCAACATTGGAAACAGGTTTTTGTAGAACTCACATAAG	5097
QY	4957	AACAAGAGGAGGAAGATAAATAGTTGTGGAACTTTTAAACAGACACACTTTGTGGGAAAGCAA	5016
DB	5098	AACAAGAGGAGGAAGATAAATAGTTGTGGAACTTTTAAACAGACACACTTTGTGGGAAAGCAA	5157
QY	5017	CACAAGAAAGATAGAGTAGTATTTTGAATTTGTATTTGTTGGTCTGCTTACTTGGGACTTCT	5076
DB	5158	CACAAGAAAGATAGAGTAGTATTTTGAATTTGTATTTGTTGGTCTGCTTACTTGGGACTTCT	5217
QY	5077	TTCTTTTTCCTAAATTTTAACTTTGGTTTAAAAAGTTTTT	5117
DB	5218	TTCTTTTTCCTAAATTTTAACTTTGGTTTAAAAAGTTTTT	5258
RESULT 4			
US-09-971-121-1			
; Sequence 1, Application US/09971121			
; Patent No. US20020111477A1			
; GENERAL INFORMATION:			
; APPLICANT: Hu, Yi			
; APPLICANT: Nepomnichy, Boris			
; TITLE OF INVENTION: No. US20020111477A1el Human Transporter Proteins and Pol			
; TITLE OF INVENTION: Same			
; FILE REFERENCE: LEX-0250-USA			
; CURRENT APPLICATION NUMBER: US/09/971,121			
; CURRENT FILING DATE: 2001-10-04			
; PRIOR APPLICATION NUMBER: US 60/239,629			
; PRIOR FILING DATE: 2000-10-10			
; NUMBER OF SEQ ID NOS: 5			
; SOFTWARE: Fast-SEQ for Windows Version 4.0			
; SEQ ID NO 1			
; LENGTH: 4929			
; TYPE: DNA			
; ORGANISM: homo sapiens			
US-09-971-121-1			
Query Match 83.5%; Score 4572; DB 9; Length 4929;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 4929; Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
QY	114	ATGTCCACTGCAATTAGGAGGTTAGGAGTTTGAGACAGACAGACAGACACTTCTACTGAAG	173
DB	1	ATGTCCACTGCAATTAGGAGGTTAGGAGTTTGAGACAGACAGACAGACACTTCTACTGAAG	60
QY	174	AATTACTTAATTAATGCAAGACCAAAAAGAGTAGTGTTCAGAAATCTTTTCCACTA	233
DB	61	AATTACTTAATTAATGCAAGACCAAAAAGAGTAGTGTTCAGAAATCTTTTCCACTA	120
QY	234	TTTTTTTTTATTTGGTTAATTAATTAAGCATGATGCATCCAAATGAAGAAATGAAGAA	293
DB	121	TTTTTTTTTATTTGGTTAATTAATTAAGCATGATGCATCCAAATGAAGAAATGAAGAA	180
QY	294	GTGCCTAAATAGAACTCAATCCTATGGACAAGTTTACTCTTTCTTAATCTTAATTTCTTGA	353
DB	181	GTGCCTAAATAGAACTCAATCCTATGGACAAGTTTACTCTTTCTTAATCTTAATTTCTTGA	240
QY	354	TATACTCCAGTGACTAATATTACAGCAGGATCATGCGAAGAGTGCTACTGATCATCTA	413
DB	241	TATACTCCAGTGACTAATATTACAGCAGGATCATGCGAAGAGTGCTACTGATCATCTA	300
QY	414	CCTGATGTCAATTAATCTGAAGAAATACAAATGAAAAAGAAATGTTAAATCCAGTCTC	473
DB	301	CCTGATGTCAATTAATCTGAAGAAATACAAATGAAAAAGAAATGTTAAATCCAGTCTC	360
QY	474	TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAGAATCCATGCTCCATGAACTTCGT	533

Db 361 TCATAGCGAGCAACTTTCTAGGTGGTGTCTTCAAGAGCTCCATGCTCTATGAACCTTCGT 420  
Qy 534 TTTTTCCTGATATGATCCAGTATCTTCTATTTATATGATTCAGAGCTGGCTGTCA 593  
Db 421 TTTTTCCTGATATGATCCAGTATCTTCTATTTATATGATTCAGAGCTGGCTGTCA 480  
Qy 594 AATCATGTGAGGCTGCTCAGTACTGCTCCAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 653  
Db 481 AATCATGTGAGGCTGCTCAGTACTGCTCCAGTCTTCAAGTCTTCAAGTCTTCAAGTCTTCA 540  
Qy 654 GATGCTGCCATTATACAGTTGAAGCAAAATGTTCTCTTTGGAAGGAGCTGGAGTCAACT 713  
Db 541 GATGCTGCCATTATACAGTTGAAGCAAAATGTTCTCTTTGGAAGGAGCTGGAGTCAACT 600  
Qy 714 AAGCTGTTATATGGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGGAGTA 773  
Db 601 AAGCTGTTATATGGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGGAGTA 660  
Qy 774 ATTTTAATATACCTAGTTATAGCATTTTCAACCTTTGGATACCTTTTGGCAATTCATATC 833  
Db 661 ATTTTAATATACCTAGTTATAGCATTTTCAACCTTTGGATACCTTTTGGCAATTCATATC 720  
Qy 834 GTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAAGGGACTTCATGATACT 893  
Db 721 GTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAAGGGACTTCATGATACT 780  
Qy 894 GCCTTTTGGCTTTCCTGGGCTTCTCTATATACAGTTTAAATTTTCTTATGCTCCCTCTT 953  
Db 781 GCCTTTTGGCTTTCCTGGGCTTCTCTATATACAGTTTAAATTTTCTTATGCTCCCTCTT 840  
Qy 954 ATGGCAGTCAATTTGGCAGACTCTTTGTTATTTCTTCAAGTAGCAGCATTTGTATATT 1013  
Db 841 ATGGCAGTCAATTTGGCAGACTCTTTGTTATTTCTTCAAGTAGCAGCATTTGTATATT 900  
Qy 1014 CTGCTTTTTCCTTATGGAATATCATCTGTATTTTGTCTTAAATGCTGACACCTCTT 1073  
Db 901 CTGCTTTTTCCTTATGGAATATCATCTGTATTTTGTCTTAAATGCTGACACCTCTT 960  
Qy 1074 TTTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTGTTTGTGCTTTTGGATTT 1133  
Db 961 TTTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTGTTTGTGCTTTTGGATTT 1020  
Qy 1134 ATTGGCTTATGATTAATCTCATAGAAAGTTTCCCAATCGTTAGTGTGCTTTTCAGT 1193  
Db 1021 ATTGGCTTATGATTAATCTCATAGAAAGTTTCCCAATCGTTAGTGTGCTTTTCAGT 1080  
Qy 1194 CTTTCTGTCACTGTACTTTTGTGATGTTATGACAGGTCATGCAATTTAGAAATTT 1253  
Db 1081 CTTTCTGTCACTGTACTTTTGTGATGTTATGACAGGTCATGCAATTTAGAAATTT 1140  
Qy 1254 AATGAGGTGCTTCAATTTCAATTTGACTGAGGCCCATATCCTCTAATTTATACAAAT 1313  
Db 1141 AATGAGGTGCTTCAATTTCAATTTGACTGAGGCCCATATCCTCTAATTTATACAAAT 1200  
Qy 1314 ATCATGCTCACACTTAATAGTATATCTATGTCCTTTTGGCTGTCTATCTTTGATCAAGTC 1373  
Db 1201 ATCATGCTCACACTTAATAGTATATCTATGTCCTTTTGGCTGTCTATCTTTGATCAAGTC 1260  
Qy 1374 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAGGCCCTCATATTGG 1433  
Db 1261 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAGGCCCTCATATTGG 1320  
Qy 1434 TCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCAATGTTAATGGAAATATTAGT 1493  
Db 1321 TCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCAATGTTAATGGAAATATTAGT 1380  
Qy 1494 TTTAGTGAATTTATGAGCCAGTTTCTCAGAAATTTGTAGGAAAGAGCCATAGAAAT 1553  
Db 1381 TTTAGTGAATTTATGAGCCAGTTTCTCAGAAATTTGTAGGAAAGAGCCATAGAAAT 1440  
Qy 1554 AGTGGTATTACAGAGCATACAGAAAGAGGCTGAAATGTGGGCTTTTCAGAAATTTG 1613  
Db 1441 AGTGGTATTACAGAGCATACAGAAAGAGGCTGAAATGTGGGCTTTTCAGAAATTTG 1500

Qy 1614 TCATTTGACATATATCAGGGTCAGATTACTGCTTACTTTGGCCACAGTCGAAACAGAAAG 1673  
Db 1501 TCATTTGACATATATCAGGGTCAGATTACTGCTTACTTTGGCCACAGTCGAAACAGAAAG 1560  
Qy 1674 AGTACATTTGATGAATATCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1733  
Db 1561 AGTACATTTGATGAATATCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1620  
Qy 1734 TATGGACACAGAGCTCTCAGAAATAGATAAATGTTTGAAGCAAGAAAAATGATTTGCATT 1793  
Db 1621 TATGGACACAGAGCTCTCAGAAATAGATAAATGTTTGAAGCAAGAAAAATGATTTGCATT 1680  
Qy 1794 TGTCCACAGTTAGATATACATTTTGTGATTTTGCAGTAGTAGAAGAAATTTATCAATTTG 1853  
Db 1681 TGTCCACAGTTAGATATACATTTTGTGATTTTGCAGTAGTAGAAGAAATTTATCAATTTG 1740  
Qy 1854 GCTTCAATCAAGGGATACAGCCAAATATAATAAACAAGAGTGCGAAGGTTTACTATA 1913  
Db 1741 GCTTCAATCAAGGGATACAGCCAAATATAATAAACAAGAGTGCGAAGGTTTACTATA 1800  
Qy 1914 GATTTAGACATGACAGACTATCAAGATAACCAAGCTTAAAAATTAAGTGTGTGCTCAAAA 1973  
Db 1801 GATTTAGACATGACAGACTATCAAGATAACCAAGCTTAAAAATTAAGTGTGTGCTCAAAA 1860  
Qy 1974 AGAAAGCTGTCAATTAGGAATGCTGTTTGGGAACCCAAAGATACCTGCTGTAGATGAA 2033  
Db 1861 AGAAAGCTGTCAATTAGGAATGCTGTTTGGGAACCCAAAGATACCTGCTGTAGATGAA 1920  
Qy 2034 CCAACAGCTGGATGACCCCTGCTCGACATATTTGTAATGTAATCTTTTAAAAATCAGA 2093  
Db 1921 CCAACAGCTGGATGACCCCTGCTCGACATATTTGTAATGTAATCTTTTAAAAATCAGA 1980  
Qy 2094 AAAGCCAAATCGGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTTGA 2153  
Db 1981 AAAGCCAAATCGGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTTGA 2040  
Qy 2154 GATAGAAAGCTGTGATATACAGAAATGCTGAAATGTTGGTTCCTTCAATGTTCTCTC 2213  
Db 2041 GATAGAAAGCTGTGATATACAGAAATGCTGAAATGTTGGTTCCTTCAATGTTCTCTC 2100  
Qy 2214 AAAGTAAATTGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGTGCCACA 2273  
Db 2101 AAAGTAAATTGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGTGCCACA 2160  
Qy 2274 GAATCTCTTCTTCACTGGTTAAAACAAATATACCTGGAGCTACTTTTATTAACAAGAAAT 2333  
Db 2161 GAATCTCTTCTTCACTGGTTAAAACAAATATACCTGGAGCTACTTTTATTAACAAGAAAT 2220  
Qy 2334 GACCAACAACTTGTGTATAGCTTTCCTTTCAAGGACATGGAACAAATTTTTCAGGTTTGT 2393  
Db 2221 GACCAACAACTTGTGTATAGCTTTCCTTTCAAGGACATGGAACAAATTTTTCAGGTTTGT 2280  
Qy 2394 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCAATTTTATGTTGGTTCCTATGAGACT 2453  
Db 2281 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCAATTTTATGTTGGGTTCCTATGAGACT 2340  
Qy 2454 TTTGGAGAGCTATTTTAAAGCTTAGAAGTTGAAGCAAAATTTGAACCAAGCAATATAGT 2513  
Db 2341 TTTGGAGAGCTATTTTAAAGCTTAGAAGTTGAAGCAAAATTTGAACCAAGCAATATAGT 2400  
Qy 2514 GTATTTTACTCAGCAGCCACTGGAGGAGAAATGGAATTTCAAAATCTTTTATGAAATGGA 2573  
Db 2401 GTATTTTACTCAGCAGCCACTGGAGGAGAAATGGAATTTCAAAATCTTTTATGAAATGGA 2460  
Qy 2574 CAGAGCTTACTTATTTCTTGTGAACCAAGGCTTCTCTAGTGAGCACTGAGCCCTTTGG 2633  
Db 2461 CAGAGCTTACTTATTTCTTGTGAACCAAGGCTTCTCTAGTGAGCACTGAGCCCTTTGG 2520  
Qy 2634 AAACAAACAGATGTATCAATAGCAAAAGTTTCATTTTCTTACCTTGAACCTGAAAGTAA 2693  
Db 2521 AAACAAACAGATGTATACATAGCAAAAGTTTCATTTTCTTACCTTGAACCTGAAAGTAA 2580



QY 2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2753  
 Db 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2640  
 QY 2754 TTGGTTTCATCCTTTTAAATAATGCTGGTTCCTCAATCAAACTTGTTCAGACTTATAT 2813  
 Db 2641 TTGGTTTCATCCTTTTAAATAATGCTGGTTCCTCAATCAAACTTGTTCAGACTTATAT 2700  
 QY 2814 TTTCATAAAACCTGGAGACCAACACATATAATAACAAAGTCTGCTTCTTCAAAATCT 2873  
 Db 2701 TTTCATAAAACCTGGAGACCAACACATATAATAACAAAGTCTGCTTCTTCAAAATCT 2760  
 QY 2874 GGTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAAGCCAGAACATATAATGGTGACG 2933  
 Db 2761 GGTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAAGCCAGAACATATAATGGTGACG 2820  
 QY 2934 ATGATTAATGACAGTGACTATGATCCGTGGTCCCATAGTGGCTTTAAATGTGATG 2993  
 Db 2821 ATGATTAATGACAGTGACTATGATCCGTGGTCCCATAGTGGCTTTAAATGTGATG 2880  
 QY 2994 CATTTCAGAAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTACTATGTTTTTCTTTA 3053  
 Db 2881 CATTTCAGAAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTACTATGTTTTTCTTTA 2940  
 QY 3054 CCTATATTAGTAATATCATTAAGTAATCTTATCTTATCATTTAAATGTGACTGAAACC 3113  
 Db 2941 CCTATATTAGTAATATCATTAAGTAATCTTATCTTATCATTTAAATGTGACTGAAACC 3000  
 QY 3114 ATCCAGATCTGGAGTACCCCACTTCTTTCAGAAAATTAAGTATAGTTTTTAAATTTGAG 3173  
 Db 3001 ATCCAGATCTGGAGTACCCCACTTCTTTCAGAAAATTAAGTATAGTTTTTAAATTTGAG 3060  
 QY 3174 CTGTATTTTCAGAGCTTTCCTGGAAATCATTTGTACTGCAATGCCACCTTATCTTGGCC 3233  
 Db 3061 CTGTATTTTCAGAGCTTTCCTGGAAATCATTTGTACTGCAATGCCACCTTATCTTGGCC 3120  
 QY 3234 ATGGAAAATGACAGAAATCATTAAGATCAAAGCTTATCTCAACTTAACTTTTCAGGCTTT 3293  
 Db 3121 ATGGAAAATGACAGAAATCATTAAGATCAAAGCTTATCTCAACTTAACTTTTCAGGCTTT 3180  
 QY 3294 TTGCCATCTGCATATTTGATTTGGACAGCTGTTGTGATATCCCTTATTTTATCATTT 3353  
 Db 3181 TTGCCATCTGCATATTTGATTTGGACAGCTGTTGTGATATCCCTTATTTTATCATTT 3240  
 QY 3354 CTATTTTTCAGTGGAGCTTATGGCAATTTTCATTTGATATGATTTTATCTGTTGTA 3413  
 Db 3241 CTATTTTTCAGTGGAGCTTATGGCAATTTTCATTTGATATGATTTTATCTGTTGTA 3300  
 QY 3414 AAGTTCTCTGCTGGTTTTTGGCTTATTTGGTTATGTTTCCATCAGTTATTTCTGTTCACT 3473  
 Db 3301 AAGTTCTCTGCTGGTTTTTGGCTTATTTGGTTATGTTTCCATCAGTTATTTCTGTTCACT 3360  
 QY 3474 TATATTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTGGTCAATTTATC 3533  
 Db 3361 TATATTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTGGTCAATTTATC 3420  
 QY 3534 TATTCTGCGGAGCGTGGCTTGTATTTGCAATCACTGAATAACTTTCTTTTATGGGATAC 3593  
 Db 3421 TATTCTGCGGAGCGTGGCTTGTATTTGCAATCACTGAATAACTTTCTTTTATGGGATAC 3480  
 QY 3594 ACAATGCAACTATTTCTTCATATGCTTTTGTATATCATTCATTCOAATCTATPCCACTTTA 3653  
 Db 3481 ACAATGCAACTATTTCTTCATATGCTTTTGTATATCATTCATTCOAATCTATPCCACTTTA 3540  
 QY 3654 GGTTCCTCTGATTTCTTTTCATAAAGATTTCTTTGGAGAAATGTACGAAAATTTGGACACC 3713  
 Db 3541 GGTTCCTCTGATTTCTTTTCATAAAGATTTCTTTGGAGAAATGTACGAAAATTTGGACACC 3600  
 QY 3714 TATAATCCATGGGATAGCTTTTCAGTAGCTGTTATATCGCCTTACCTGCAAGTGTGACTG 3773  
 Db 3601 TATAATCCATGGGATAGCTTTTCAGTAGCTGTTATATCGCCTTACCTGCAAGTGTGACTG 3660  
 QY 3774 TGGATTTTCTTCTTACAACTATGAGAAAATATATGGAGCGAGATCAATAAGAAAGAT 3833

Db 3661 TGGATTTTCTTCTTACAACTATGAGAAAATATATGGAGCGAGATCAATAAGAAAGAT 3720  
 QY 3834 CCCTTTTTCAGAAAACCTTTTCAACGAAAGTCTTAAATAATAGGAAGCTTCCAGAACCCAGAC 3893  
 Db 3721 CCCTTTTTCAGAAAACCTTTTCAACGAAAGTCTTAAATAATAGGAAGCTTCCAGAACCCAGAC 3780  
 QY 3894 AATGAGGATGAAGATGAGATGTCAAAGCTGAAAGCTTAAAGCTCAAGAGCTGATGGGT 3953  
 Db 3781 AATGAGGATGAAGATGAGATGTCAAAGCTGAAAGCTTAAAGCTCAAGAGCTGATGGGT 3840  
 QY 3954 TGGCAGTGTGTGAGGAGAAACCATCCATTATGGTCAGCAATTTGTCATATAAGAAATATGAT 4013  
 Db 3841 TGGCAGTGTGTGAGGAGAAACCATCCATTATGGTCAGCAATTTGTCATATAAGAAATATGAT 3900  
 QY 4014 GACAGAAAGATTTTCTTCTTTCAGAAAATGAAAGCTTAAAGCTGAAAGCTGATGGGT 4073  
 Db 3901 GACAGAAAGATTTTCTTCTTTCAGAAAATGAAAGCTTAAAGCTGAAAGCTGATGGGT 3960  
 QY 4074 TTCTGTGTGAAAAGAGAGAGATCTTTCAGCAATTTGGGTCCAAATGGTCTGCGAAAAGC 4133  
 Db 3961 TTCTGTGTGAAAAGAGAGAGATCTTTCAGCAATTTGGGTCCAAATGGTCTGCGAAAAGC 4020  
 QY 4134 ACAATTTTAAATATCTGTTGGTGATATGTAACCAACTTCAGGCCAGGTATTTTTAGGA 4193  
 Db 4021 ACAATTTTAAATATCTGTTGGTGATATGTAACCAACTTCAGGCCAGGTATTTTTAGGA 4080  
 QY 4194 GATTATCTTCAGAGCAAGTGAAGATGATGATTCATCTGAAGTGTATGGTTACTGCTCT 4253  
 Db 4081 GATTATCTTCAGAGCAAGTGAAGATGATGATTCATCTGAAGTGTATGGTTACTGCTCT 4140  
 QY 4254 CAGATAAAACCTTTTGGCCAGATACATTTGCAAGAAACATTTTGAATTTATGGAGCT 4313  
 Db 4141 CAGATAAAACCTTTTGGCCAGATACATTTGCAAGAAACATTTTGAATTTATGGAGCT 4200  
 QY 4314 GTCAGAGGAATGAGTGCAGAGTGCATGAAAGTGCATTAAGTGCATACATGCACCT 4373  
 Db 4201 GTCAGAGGAATGAGTGCAGAGTGCATGAAAGTGCATTAAGTGCATACATGCACCT 4260  
 QY 4374 GATTTTAAAGAAACATCTTCAGAAAGCTGTAAAGAAACCTTCGAGGAATCAAAAGAAAG 4433  
 Db 4261 GATTTTAAAGAAACATCTTCAGAAAGCTGTAAAGAAACCTTCGAGGAATCAAAAGAAAG 4320  
 QY 4434 TTGTGTTTTTCTTAAAGTATGCTAGGATCTCCTCAGATTTACTTTGCTAGATGAACCATCT 4493  
 Db 4321 TTGTGTTTTTCTTAAAGTATGCTAGGATCTCCTCAGATTTACTTTGCTAGATGAACCATCT 4380  
 QY 4494 ACAGGTATGGATCCCAAGCCAAACAGACATGTCGCGAGCAATTCGAACTGCATTTAAA 4553  
 Db 4381 ACAGGTATGGATCCCAAGCCAAACAGACATGTCGCGAGCAATTCGAACTGCATTTAAA 4440  
 QY 4554 AACAGAAACCGGCTGCTTATCTGACCACTCATATATGGAGGAGGAGGCTGCTGCT 4613  
 Db 4441 AACAGAAACCGGCTGCTTATCTGACCACTCATATATGGAGGAGGAGGCTGCTGCT 4500  
 QY 4614 GATCAGATGATCATCATGTTGCTGCGCAAGTTCAGATGTCGGAACAGTCAACATCTA 4673  
 Db 4501 GATCAGATGATCATCATGTTGCTGCGCAAGTTCAGATGTCGGAACAGTCAACATCTA 4560  
 QY 4674 AAGAGTAAATTTGGAAAAGGCTTCTTTTGGAAAATTAATTTGAAGGATCTGGATGAAAAC 4733  
 Db 4561 AAGAGTAAATTTGGAAAAGGCTTCTTTTGGAAAATTAATTTGAAGGATCTGGATGAAAAC 4620  
 QY 4734 CTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATTCGAAGCGCTCAG 4793  
 Db 4621 CTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATTCGAAGCGCTCAG 4680  
 QY 4794 GAAAGTTTTTCTTCTTATTTTGGCTTATATAATTCCTAAGGAAGATGTTTCAGTCCCTTTCA 4853  
 Db 4681 GAAAGTTTTTCTTCTTATTTTGGCTTATATAATTCCTAAGGAAGATGTTTCAGTCCCTTTCA 4740  
 QY 4854 CAACTCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCTTTCGCAATGGAAGATATAGCTTT 4913

Db 4741 CAATCTTTTAAAGCTGAAGAGCTAAACATGCTTTTGGCAATTGAAGATAATAGCTTT 4800  
 QY 4914 TCTAAGCAACATTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGGAAGAT 4973  
 Db 4801 TCTAAGCAACATTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGGAAGAT 4860  
 QY 4974 AATAGTTGTGGAACCTTTAAACAGCACCTTTTGGTGGGAACGAACAACAGAGAGATAGAGTA 5033  
 Db 4861 AATAGTTGTGGAACCTTTAAACAGCACCTTTTGGTGGGAACGAACAACAGAGAGATAGAGTA 4920  
 QY 5034 GTATTTTGA 5042  
 Db 4921 GTATTTTGA 4929

RESULT 5  
 US-09-971-121-3  
 ; Sequence 3, Application US/09971121  
 ; Patent No. US20020111477A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hu, Yi  
 ; TITLE OF INVENTION: Same  
 ; TITLE OF INVENTION: Same  
 ; FILE REFERENCE: LEX-0250-USA  
 ; CURRENT APPLICATION NUMBER: US/09/971,121  
 ; PRIOR FILING DATE: 2001-10-04  
 ; PRIOR APPLICATION NUMBER: US 60/239,629  
 ; NUMBER OF SEQ ID NOS: 5  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 3  
 ; LENGTH: 4785  
 ; TYPE: DNA  
 ; ORGANISM: homo sapiens  
 US-09-971-121-3

Query Match 80.5%; Score 4408; DB 9; Length 4785;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 4758; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 114 ATGTCCTCACTCAATTTAGGAGGTAGAGCTTTGGAGACAGACAGACCACTTCTACTGAAG 173  
 Db 1 ATGTCCTCACTCAATTTAGGAGGTAGAGCTTTGGAGACAGACCACTTCTACTGAAG 60

QY 174 AATTAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 233  
 Db 61 AATTAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 120

QY 234 TTTTCTTTTATTTTGTGTTAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 293  
 Db 121 TTTTCTTTTATTTTGGTTAATTAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 180

QY 294 GTGCTAATATAGAACTCAATCCTATGGAACAGTTTACTCTTTCTAATCTAATTTCTTTGA 353  
 Db 181 GTGCTAATATAGAACTCAATCCTATGGAACAGTTTACTCTTTCTAATCTAATTTCTTTGA 240

QY 354 TATACCTCAGTCACTAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 413  
 Db 241 TATACCTCAGTCACTAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 300

QY 414 CTGATGTCAATAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 473  
 Db 301 CTGATGTCAATAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 360

QY 474 TCTAAGCCGAGCACTTTGTAGGTGTGGTTTCAAGACTCCATGCTCTATGAACTTCGT 533  
 Db 361 TCTAAGCCGAGCACTTTGTAGGTGTGGTTTCAAGACTCCATGCTCTATGAACTTCGT 420

QY 534 TTTTCTCTGATATGATTCAGTATCTTCTATTTATATGATTCAGGAGTGGCTGTTC 593  
 Db 421 TTTTCTCTGATATGATTCAGTATCTTCTATTTATATGATTCAGGAGTGGCTGTTC 480

QY 594 AATCAATGTGAGCTCTCTCAGTACTCGTCTCAGGTTTCAAGTTTCAAGCATCATATA 653  
 Db 481 AATCAATGTGAGCTCTCTCAGTACTCGTCTCAGGTTTCAAGTTTCAAGCATCATATA 540  
 QY 654 GATGCTGCCAATATATACAGTTGAAGCAATGTTTCTCTTTGGAAGAGCTGGAGTCAACT 713  
 Db 541 GATGCTGCCAATATATACAGTTGAAGCAATGTTTCTCTTTGGAAGAGCTGGAGTCAACT 600  
 QY 714 AAGCTGTTTATATAGGAGAAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGTA 773  
 Db 601 AAGCTGTTTATATAGGAGAAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGTA 660  
 QY 774 ATTTTAATATACCTAGTATAGCAATTTTCCCTTTTGGATCTTTTGGCAATTCATATC 833  
 Db 661 ATTTTAATATACCTAGTATAGCAATTTTCCCTTTTGGATCTTTTGGCAATTCATATC 720  
 QY 834 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGACTT CATGATAT 893  
 Db 721 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGACTT CATGATAT 780  
 QY 894 GCCTTTTGGCTTTCCCTGGGTCTCTCTATATACAAGTTTAAATTTTCTTATGTCCTTCT 953  
 Db 781 GCCTTTTGGCTTTCCCTGGGTCTCTCTATATACAAGTTTAAATTTTCTTATGTCCTTCT 840  
 QY 954 ATGGCAGTCAATTCGACAGCTTCTTTTGTATTTTCTCAAAGTAGCAGCAATTTGTATTT 1013  
 Db 841 ATGGCAGTCAATTCGACAGCTTCTTTTGTATTTTCTCAAAGTAGCAGCAATTTGTATTT 900  
 QY 1014 CTGCTTTTCTTTTATGGAATATCATCTGATTTTCTTAAATGCTGACCTCTT 1073  
 Db 901 CTGCTTTTCTTTTATGGAATATCATCTGATTTTCTTAAATGCTGACCTCTT 960  
 QY 1074 TTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTACTGTGCTTTTGGATTT 1133  
 Db 961 TTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTACTGTGCTTTTGGATTT 1020  
 QY 1134 ATGGCTTTATGATTAATCTCTATAGAAAGTTTCCAAATCGTTAGTGGCTTTTCAGT 1193  
 Db 1021 ATGGCTTTATGATTAATCTCTATAGAAAGTTTCCAAATCGTTAGTGGCTTTTCAGT 1080  
 QY 1194 CTTTCTGTCTCACTGACTTTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1253  
 Db 1081 CTTTCTGTCTCACTGACTTTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 1140  
 QY 1254 AATGAAGTGTCTCAATTTCAAAATTTGACTGAGGCCCATATCTCTTAATTTACAAT 1313  
 Db 1141 AATGAAGTGTCTCAATTTCAAAATTTGACTGAGGCCCATATCTCTTAATTTACAAT 1200  
 QY 1314 ATCATGCTCACACTTAAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTCAATGTC 1373  
 Db 1201 ATCATGCTCACACTTAAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTCAATGTC 1260  
 QY 1374 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATATTTTCTGAAGCTTCATATGG 1433  
 Db 1261 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATATTTTCTGAAGCTTCATATGG 1320  
 QY 1434 TCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCAATGTTAAATGGAATATTAGT 1493  
 Db 1321 TCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCAATGTTAAATGGAATATTAGT 1380  
 QY 1494 TTTAGTGAATTTATGAGCAGCTTTCTTCAAGAAATTTAGGAAAGAGCCATAGAAT 1553  
 Db 1381 TTTAGTGAATTTATGAGCAGCTTTCTTCAAGAAATTTAGGAAAGAGCCATAGAAT 1440  
 QY 1554 AGTGTNTTCAAGACATACAGAGAGGAGGTTGAATTTGAGGCTTTGAGAAATTTG 1613  
 Db 1441 AGTGTNTTCAAGACATACAGAGAGGAGGTTGAATTTGAGGCTTTGAGAAATTTG 1500  
 QY 1614 TCATTTGACATATATGAGGCTCAGATTTACTGCTTTTCTTGGCCACAGTGGAAACAGGAAG 1673  
 Db 1501 TCATTTGACATATATGAGGCTCAGATTTACTGCTTTTCTTGGCCACAGTGGAAACAGGAAG 1560  
 QY 1674 AGTACATTCATGAATATCTTTTGTGGACTCTGCCACACCTTCTGTGAGGTTTGCATCTATA 1733



1561 AGTACATTGATGAATATTTCTTTGGGACTCTGCCACCTTCTGATGGTTTGGCATCTATA 1620  
1734 TATGAGACACAGAGTCTCGAATAATAGATGAATGTTTGAAGCAGAAAAATGATTGGCATT 1793  
1621 TATGAGACACAGAGTCTCGAATAATAGATGAATGTTTGAAGCAGAAAAATGATTGGCATT 1680  
1794 TGTCCACAGTTAGATATACATCTTTGATGCTTTTGGAGTGTGACGTAGAGAAAAATTTATCAATTTTG 1853  
1681 TGTCCACAGTTAGATATACATCTTTGATGCTTTTGGAGTGTGACGTAGAGAAAAATTTATCAATTTTG 1740  
1854 GCTTCAATCAAGGGATACCGACCAACAAATATAATACAAAGAGTGCAGAGGTTTACTA 1913  
1741 GCTTCAATCAAGGGATACCGACCAACAAATATAATACAAAGAGTGCAGAGGTTTACTA 1800  
1914 GATTTAGACATGACAGTATCAAGATTAACCAAGCTAAATAATTAAGTGTGTGTCAAAAA 1973  
1801 GATTTAGACATGACAGTATCAAGATTAACCAAGCTAAATAATTAAGTGTGTGTCAAAAA 1860  
1974 AGAAGCTGTCTATTAGGAATTTGCTGTTCTTGGGAACCAAGATATCTGCTGTAGATGAA 2033  
1861 AGAAGCTGTCTATTAGGAATTTGCTGTTCTTGGGAACCAAGATATCTGCTGTAGATGAA 1920  
2034 CCAACAGCTGAATGGACCCCTGTTCTCGACATATTTGTATGGAATCTTTTAAAAATACAGA 2093  
1921 CCAACAGCTGAATGGACCCCTGTTCTCGACATATTTGTATGGAATCTTTTAAAAATACAGA 1980  
2094 AAAGCCAATCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTTGCA 2153  
1981 AAAGCCAATCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTTGCA 2040  
2154 GATAGGAAGCTGTGATATCAAAAGGAATGCTGGAATGTGTGTTCTTCAATGTTCCTC 2213  
2041 GATAGGAAGCTGTGATATCAAAAGGAATGCTGGAATGTGTGTTCTTCAATGTTCCTC 2100  
2214 AAAAGTAATCGGGGATCGGCTACCGCTGACGATGTACATAGCAAAATTTTGGCCACA 2273  
2101 AAAAGTAATCGGGGATCGGCTACCGCTGACGATGTACATAGCAAAATTTTGGCCACA 2160  
2274 GAATCTCTTTCTTCACTGGTTTAAACAAATATACCTGGAGCTACTTTATTAACACAGAAT 2333  
2161 GAATCTCTTTCTTCACTGGTTTAAACAAATATACCTGGAGCTACTTTATTAACACAGAAT 2220  
2334 GACCAACAACTTGTGTATAGTTGCTTTCAGAGCATGACAAAAATTTTCAAGTTTGT 2393  
2221 GACCAACAACTTGTGTATAGTTGCTTTCAGAGCATGACAAAAATTTTCAAGTTTGT 2280  
2394 TCTGCCCTTAGACAGTCAATCAAAATTTGGGTGTCTATTTCTTATGCTGTTTCCATGACGACT 2453  
2281 TCTGCCCTTAGACAGTCAATCAAAATTTGGGTGGCATTTCTTATGGGTTCATGACGACT 2340  
2454 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATAGT 2513  
2341 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATAGT 2400  
2514 GTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTTCAAAATCTTTTGTATGAATGGAA 2573  
2401 GTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTTCAAAATCTTTTGTATGAATGGAA 2460  
2574 CAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTTCTTAGTGAGCACCATGAGCCCTTGG 2633  
2461 CAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTKCTTAGTGAGCACCATGAGCCCTTGG 2520  
2634 AAAACAAGATGTATACAAATAGCAAGTTTCAATTTCTTTTACCTTGAACGTGAAGTAAA 2693  
2521 AAAACAAGATGTATACAAATAGCAAGTTTCAATTTCTTTTACCTTGAACGTGAAGTAAA 2580  
2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTTCAGATTTTATGTTT 2753  
2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTTCAGATTTTATGTTT 2640  
2754 TTGGTTTCAATCCTTTTAAATGCTGTGGTTTCCCATGAACTTGTTCAGACCTTATAT 2813

2641 TTGGTTTCAATCCTCTTTTAAAAATGCTGTGGTTCCCATCAAACTTGTTCAGACCTTATAT 2700  
2814 TTTCTAAAAACCTGGAGACAAACACATATAATACAAACAAAGCTGTGCTTCTTCAAAATTTCT 2873  
2701 TTTCTAAAAACCTGGAGACAAACACATATAATACAAACAAAGCTGTGCTTCTTCAAAATTTCT 2760  
2874 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAGGCAAGACATAATGTTGACG 2933  
2761 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAGGCAAGACATAATGTTGACG 2820  
2934 ATGATTTAATGACAGTACTATGATCCGTGGCTCCCATAGTCGGCTTTAAATGTTGATG 2993  
2821 ATGATTTAATGACAGTACTATGATCCGTGGCTCCCATAGTCGGCTTTAAATGTTGATG 2880  
2994 CATTCAGAAAAAGCACTATGTTTTCAGAGCTGTTTTCACAGTACTATGTTTATTTCTTTA 3053  
2881 CATTCAGAAAAAGCACTATGTTTTCAGAGCTGTTTTCACAGTACTATGTTTATTTCTTTA 2940  
3054 CTTATATTAGTCAATATCATTAGTAACTACTATCTTTATCATTTAAATGTTGACTCAAAAC 3113  
2941 CTTATATTAGTCAATATCATTAGTAACTACTATCTTTATCATTTAAATGTTGACTCAAAAC 3000  
3114 ATCCAGATCTGAGTACCCCATCTTTTCAGAAATTAATGATATAGTTTAAAAATTTGAG 3173  
3001 ATCCAGATCTGAGTACCCCATCTTTTCAGAAATTAATGATATAGTTTAAAAATTTGAG 3060  
3174 CTGTATTTTCAAGCAGCTTTGCTTGGAACTATTGTTTACTGCAATGCACTTACTTTTGGC 3233  
3061 CTGTATTTTCAAGCAGCTTTGCTTGGAACTATTGTTTACTGCAATGCACTTACTTTTGGC 3120  
3234 ATGGAATATGACAGAAATCAATAAGATCAAAAGCTTATATCAACTTAAATTTTCAAGTCTT 3293  
3121 ATGGAATATGACAGAAATCAATAAGATCAAAAGCTTATATCAACTTAAATTTTCAAGTCTT 3180  
3294 TTGGCCATCTGCATATTTGGATTGGACAGCTGTTGATATGCCCTTATTTTATCTTATCAT 3353  
3181 TTGGCCATCTGCATATTTGGATTGGACAGCTGTTGATATGCCCTTATTTTATCTTATCAT 3240  
3354 CTTATTTTATGCTAGGAGCTTATTTGGCATTTTCAATTATGATTTATTTTATACTGTA 3413  
3241 CTTATTTTATGCTAGGAGCTTATTTGGCATTTTCAATTATGATTTATTTTATACTGTA 3300  
3414 AGTTTCTTCTGCTGTTTGGCTTATGCTTATGTTTATGTTTCCATCAGTTATTTCTGTTCA 3473  
3301 AGTTTCTTCTGCTGTTTGGCTTATGCTTATGTTTATGTTTCCATCAGTTATTTCTGTTCA 3360  
3474 TATATTTGCTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3533  
3361 TATATTTGCTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
3534 TATTTCTGTCGACGCTTGGCTTATGCTTATGCTTATGTTTATGTTTATGTTTATGTTATC 3593  
3421 TATTTCTGTCGACGCTTGGCTTATGCTTATGCTTATGTTTATGTTTATGTTTATGTTATC 3480  
3594 ACATTTGCAATCTATTTCTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATC 3653  
3481 ACATTTGCAATCTATTTCTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATC 3540  
3654 GGTTCCTCTGATTTCTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATC 3713  
3541 GGTTCCTCTGATTTCTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATC 3600  
3714 TATATCCATCGGATAGGCTTTCAGTAGCTGTTTATGCTTATGCTTATGCTTATGCTTATGCTTATC 3773  
3601 TATATCCATCGGATAGGCTTTCAGTAGCTGTTTATGCTTATGCTTATGCTTATGCTTATGCTTATC 3660  
3774 TGGATTTTCTCTTACAACTATGAGAAAAATATGAGGAGCAGATCAATTAAGAAAAAGAT 3833  
3661 TGGATTTTCTCTTACAACTATGAGAAAAATATGAGGAGCAGATCAATTAAGAAAAAGAT 3720  
3834 CCCTTTTTCAGAAACCTTTTCAACGAGTCTAAATAGGAGGCTTCCAGAACCCACAC 3893  
3721 CCCTTTTTCAGAAACCTTTTCAACGAGTCTAAATAGGAGGCTTCCAGAACCCACAC 3780

3894 AATGAGATGAAGATGAAGATCTCAAGCTGAAGACTAAAGGTCAAGAGACTGATGGT 3953  
 Db AATGAGATGAAGATGAAGATCTCAAGCTGAAGACTAAAGGTCAAGAGACTGATGGT 3840  
 QY TGCCAGTGTGTGAGGAGAACCCATTCATATGCTCAGCAATTTGCATAAAGATATGAT 4013  
 Db TGCCAGTGTGTGAGGAGAACCCATTCATATGCTCAGCAATTTGCATAAAGATATGAT 3900  
 QY GACAAAGAGATTTTCTTTTCAAGAAAGATGAAAGAGTGGCAACTTAAATACATCTCT 4073  
 Db GACAAAGAGATTTTCTTTTCAAGAAAGATGAAAGAGTGGCAACTTAAATACATCTCT 3960  
 QY TTCTGTGTCGAAAGAGAGATCTTAGGACTATTTGGCTCCAAATGCTGGCAAGAGC 4133  
 Db TTCTGTGTCGAAAGAGAGATCTTAGGACTATTTGGCTCCAAATGCTGGCAAGAGC 4020  
 QY ACAATTAATTAATTTCTGTTGGTGGATATGCAACCACTTCAGGCCAGGTATTTTAGGA 4193  
 Db ACAATTAATTAATTTCTGTTGGTGGATATGCAACCACTTCAGGCCAGGTATTTTAGGA 4080  
 QY GATTATTTCTTCAGACAAAGTGAAGATGATGATTCACCTGAAGTGTATGGGTACTGTCCT 4253  
 Db GATTATTTCTTCAGACAAAGTGAAGATGATGATTCACCTGAAGTGTATGGGTACTGTCCT 4140  
 QY CAGATAAAACCTTTGTCGCCAGATACATACATTCAGAGAACATTTTGAATTTTATGGAGCT 4313  
 Db CAGATAAAACCTTTGTCGCCAGATACATACATTCAGAGAACATTTTGAATTTTATGGAGCT 4200  
 QY GTCAAGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4373  
 Db GTCAAGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4260  
 QY GTCAAGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4433  
 Db GTCAAGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4320  
 QY TTGTTGTTTGTCTAAGTATGCTAGGAAATCTCTAGATTAATTTGCTAGATGAACCATCT 4493  
 Db TTGTTGTTTGTCTAAGTATGCTAGGAAATCTCTAGATTAATTTGCTAGATGAACCATCT 4380  
 QY ACAGGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4553  
 Db ACAGGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4440  
 QY AACAGAGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4613  
 Db AACAGAGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4500  
 QY GATCGAGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4673  
 Db GATCGAGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4560  
 QY AAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAGGACTGGATAGAAAC 4733  
 Db AAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAGGACTGGATAGAAAC 4620  
 QY CTAGAAGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4793  
 Db CTAGAAGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4680  
 QY GAAAGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4853  
 Db GAAAGTATGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACTGCAATTTAAA 4740  
 QY CAATCTTTTAAAGTGAAGAG 4878  
 Db CAATCTTTTAAAGTGAAGAG 4765

RESULT 6  
 US-10-090-458-1  
 ; Sequence 1, Application US/10090458

Publication No. US20020123107A1  
 GENERAL INFORMATION:  
 APPLICANT: Chen, Hongyun  
 APPLICANT: Kilinski, Ligia  
 APPLICANT: Le Bihan, Stephane  
 TITLE OF INVENTION: NOVEL ABCAS TRANSPORTER AND USES THEREOF  
 FILE REFERENCE: 100103.403  
 CURRENT APPLICATION NUMBER: US/10/090,458  
 CURRENT FILING DATE: 2002-03-01  
 NUMBER OF SEQ ID NOS: 5  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 1  
 LENGTH: 5463  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: 25, 2888, 2889  
 OTHER INFORMATION: n = A,T,C or G  
 US-10-090-458-1

Query Match 51.8%; Score 2826; DB 13; Length 5463;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2876; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCTCCGCGCCCTCGCACAGATCCNAGCTGGTCAACCCGACCTGAGTCAACAGACTCGAGC 60  
 Db 1 GCCTCCGCGCCCTCGCACAGATCCNAGCTGGTCAACCCGACCTGAGTCAACAGACTCGAGC 60  
 QY 61 GCCTCAGGCTGACAGCTCTGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCT 120  
 Db 61 GCCTCAGGCTGACAGCTCTGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCT 120  
 QY 121 CTCGAATAGGAGGTAGGAGTTGGAGACAGACAGACAGACACTTCTACTGGAAGATTAAT 180  
 Db 121 CTCGAATAGGAGGTAGGAGTTGGAGACAGACAGACAGACACTTCTACTGGAAGATTAAT 180  
 QY 181 TAATTAATGCGAGAACCAAAAGAGTAGTGTTCAGGAAATCTCTTTTCCACTATTTTTT 240  
 Db 181 TAATTAATGCGAGAACCAAAAGAGTAGTGTTCAGGAAATCTCTTTTCCACTATTTTTT 240  
 QY 241 TATTTTGGTAAATATTAATTAATTAGCATGATGCATCCCAATTAAGAAATATGAAGTGCCTA 300  
 Db 241 TATTTTGGTAAATATTAATTAATTAGCATGATGCATCCCAATTAAGAAATATGAAGTGCCTA 300  
 QY 301 ATATGAAGTCAATCTCTATGGAAGTGTCTCTTTTCTAATCTAATCTTCTTGGATATCTC 360  
 Db 301 ATATGAAGTCAATCTCTATGGAAGTGTCTCTTTTCTAATCTAATCTTCTTGGATATCTC 360  
 QY 361 CAGTGACTAATATTAAGACAGCATCATGAGAGTGTCTACTGATCATCTACCTGATG 420  
 Db 361 CAGTGACTAATATTAAGACAGCATCATGAGAGTGTCTACTGATCATCTACCTGATG 420  
 QY 421 TCATAATTAATGAGAAATATCAAAATGAAAAAGAAATGTTAAATCATCCAGTCTCTCTAAGC 480  
 Db 421 TCATAATTAATGAGAAATATCAAAATGAAAAAGAAATGTTAAATCATCCAGTCTCTCTAAGC 480  
 QY 481 CGAGCAACTTTGTAGTGTGGTTCCTCAAGAGTCCATGCTCTTCAAGTCTCTGTTTTTTC 540  
 Db 481 CGAGCAACTTTGTAGTGTGGTTCCTCAAGAGTCCATGCTCTTCAAGTCTCTGTTTTTTC 540  
 QY 541 CTGATATGATCCAGTATCTTCTATTTATGATGATTTCAAGAGTGGCTGCTTCAAAATCAT 600  
 Db 541 CTGATATGATCCAGTATCTTCTATTTATGATGATTTCAAGAGTGGCTGCTTCAAAATCAT 600  
 QY 601 GTGAGGCTCTCAGTACTGCTCTCAGTGTTCAGTGTTCAGTGTTCAGTGTTCAGTGTTCAGTGT 660  
 Db 601 GTGAGGCTCTCAGTACTGCTCTCAGTGTTCAGTGTTCAGTGTTCAGTGTTCAGTGTTCAGTGT 660  
 QY 661 CCATTATCAGTTGAAGACCAATGTTCTCTTTGGAAGAGTGGAGTCAACTAAAGCTG 720  
 Db 661 CCATTATCAGTTGAAGACCAATGTTCTCTTTGGAAGAGTGGAGTCAACTAAAGCTG 720

QY 721 TTATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTAATTTAA 780  
Db 721 TTATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTAATTTAA 780  
QY 781 TATACCTAGTTATAGCAATTTTACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG 840  
Db 781 TATACCTAGTTATAGCAATTTTACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG 840  
QY 841 AAAAAAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTTCATGATACCTGCTTTT 900  
Db 841 AAAAAAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTTCATGATACCTGCTTTT 900  
QY 901 GGCCTTCCTGGTCTCTCTATATACAAGTTTAAATTTTCTTATGTCCTTCTTATGCGCAG 960  
Db 901 GGCCTTCCTGGTCTCTCTATATACAAGTTTAAATTTTCTTATGTCCTTCTTATGCGCAG 960  
QY 961 TCATTGGCAGCAGCTCTCTGTTATTTCTCTCAAGTATGAGCATTTGTGATATTTCTGCTTT 1020  
Db 961 TCATTGGCAGCAGCTCTCTGTTATTTCTCTCAAGTATGAGCATTTGTGATATTTCTGCTTT 1020  
QY 1021 TTTTCCCTTATGGATTATCATCTGTATTTTGTCTTTAATGCTGACACCTCTTTTAAAA 1080  
Db 1021 TTTTCCCTTATGGATTATCATCTGTATTTTGTCTTTAATGCTGACACCTCTTTTAAAA 1080  
QY 1081 AATCAAAACATGTGGGAATAGTTGAAATTTTGTGTTAATGCTGCTTTTGGATTTATGGCC 1140  
Db 1081 AATCAAAACATGTGGGAATAGTTGAAATTTTGTGTTAATGCTGCTTTTGGATTTATGGCC 1140  
QY 1141 TTATGATATCTCTATAGAAAGTTTCCCAATCGTTAGTGTGCTTTTCACTGCTTTCT 1200  
Db 1141 TTATGATATCTCTATAGAAAGTTTCCCAATCGTTAGTGTGCTTTTCACTGCTTTCT 1200  
QY 1201 GTCACTGTACTTTTGTGATGTTATGTCAGAGTCAATTTAGAAATTTTAAATGAAG 1260  
Db 1201 GTCACTGTACTTTTGTGATGTTATGTCAGAGTCAATTTAGAAATTTTAAATGAAG 1260  
QY 1261 GTGCTTCAATTTTCAAAATTTGATGAGGCTTATGCTGCTTTTCACTGCTTTCT 1320  
Db 1261 GTGCTTCAATTTTCAAAATTTGATGAGGCTTATGCTGCTTTTCACTGCTTTCT 1320  
QY 1321 TCACACTTAAATAGTATATCTATGTCCTTTGGTGTCTATCTTTGATCAAGTCAATTCAG 1380  
Db 1321 TCACACTTAAATAGTATATCTATGTCCTTTGGTGTCTATCTTTGATCAAGTCAATTCAG 1380  
QY 1381 GGGAAATTTGGCTTACGGAGATCACTTTTATATTTTCTGAAGCCTTCATATTTGTCAAAGA 1440  
Db 1381 GGGAAATTTGGCTTACGGAGATCACTTTTATATTTTCTGAAGCCTTCATATTTGTCAAAGA 1440  
QY 1441 GCAAAAGAAATATGAGGAGTATCAGAGGCAATTTAATGGAATATTAGTTTATAGTG 1500  
Db 1441 GCAAAAGAAATATGAGGAGTATCAGAGGCAATTTAATGGAATATTAGTTTATAGTG 1500  
QY 1501 AAATTTATGAGCAGTTCTTCAGAAATTTGAGGAAGAACCAATGAATTTAGTGCTGA 1560  
Db 1501 AAATTTATGAGCAGTTCTTCAGAAATTTGAGGAAGAACCAATGAATTTAGTGCTGA 1560  
QY 1561 TFCAGAAGACATACAGAAAGAGGGTGAATTTGAGGCTTTGAGAAATTTGCTATTG 1620  
Db 1561 TFCAGAAGACATACAGAAAGAGGGTGAATTTGAGGCTTTGAGAAATTTGCTATTG 1620  
QY 1621 ACATATATGAGGCTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGGAAGAGTACAT 1680  
Db 1621 ACATATATGAGGCTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGGAAGAGTACAT 1680  
QY 1681 TGATGAATTTCTTTGTCGACTCTGCCCCCTCTGATGGTTTGCATCTATATATGAC 1740  
Db 1681 TGATGAATTTCTTTGTCGACTCTGCCCCCTCTGATGGTTTGCATCTATATATGAC 1740  
QY 1741 ACAGAGTCTCAGAAATAGATGAATTTGTTGAAGCAAGAAATGATTTGGCATTTTGTCCAC 1800  
Db 1741 ACAGAGTCTCAGAAATAGATGAATTTGTTGAAGCAAGAAATGATTTGGCATTTTGTCCAC 1800  
QY 1801 AGTTAGATATACACTTTGATGTTTGTGACAGTAGAAGAAATTTATCAATTTTGGCTCAA 1860

Db 1801 AGTTAGATATACACTTTGATGTTTGTGACAGTAGAAGAAATTTATCAATTTTGGCTCAA 1860  
QY 1861 TCAAGGGATACAGCCAAACAATATATAACAAGTGCAGAGGTTTTTACTAGATTTAG 1920  
Db 1861 TCAAGGGATACAGCCAAACAATATATAACAAGTGCAGAGGTTTTTACTAGATTTAG 1920  
QY 1921 ACATGAGACTATCAAGATACCAAGCTTAAAAAATTTAGTGTGCTCAAAAAGAAAGC 1980  
Db 1921 ACATGAGACTATCAAGATACCAAGCTTAAAAAATTTAGTGTGCTCAAAAAGAAAGC 1980  
QY 1981 TGTCAATTAGGAATTTGCTTTCTTTGGGAACCCAAAGATCTGCTGTAGATGAACCAACAG 2040  
Db 1981 TGTCAATTAGGAATTTGCTTTCTTTGGGAACCCAAAGATCTGCTGTAGATGAACCAACAG 2040  
QY 2041 CTGGAAATGACCCCTGTTCTCGACATATTTGTATGGAATCTTTTAAATACAGAAAGCCA 2100  
Db 2041 CTGGAAATGACCCCTGTTCTCGACATATTTGTATGGAATCTTTTAAATACAGAAAGCCA 2100  
QY 2101 ATCGGCTGACAGTGTTCAGTACTCTTTTCATGATGAAGCTGACATTTCTTGAGATAGGA 2160  
Db 2101 ATCGGCTGACAGTGTTCAGTACTCTTTTCATGATGAAGCTGACATTTCTTGAGATAGGA 2160  
QY 2161 AGCTGTGATATCAACAAGGAATGCTGAATGTTGTTCTTCAATGTTCTTCAAAAGTA 2220  
Db 2161 AGCTGTGATATCAACAAGGAATGCTGAATGTTGTTCTTCAATGTTCTTCAAAAGTA 2220  
QY 2221 AATGGGGATCGGTACCGCTGAGCATGTATAGACAAATATTTGTGCGACAGAAATCTC 2280  
Db 2221 AATGGGGATCGGTACCGCTGAGCATGTATAGACAAATATTTGTGCGACAGAAATCTC 2280  
QY 2281 TTTCTTCACTGTTTAAACAACATATACCTGGAGCTACTTTTATTAACAAGAAATGACCAAC 2340  
Db 2281 TTTCTTCACTGTTTAAACAACATATACCTGGAGCTACTTTTATTAACAAGAAATGACCAAC 2340  
QY 2341 AACTTGTGATAGCTTTCCTTTCAAGGACATGACAAATTTTTCAGTTTGTCTTCTGCC 2400  
Db 2341 AACTTGTGATAGCTTTCCTTTCAAGGACATGACAAATTTTTCAGTTTGTCTTCTGCC 2400  
QY 2401 TAGACAGTCAATCAAAATTTGGGTGTCAATTTCTATGTTGTTTCCATGACGACTTTGGAAG 2460  
Db 2401 TAGACAGTCAATCAAAATTTGGGTGTCAATTTCTATGTTGTTTCCATGACGACTTTGGAAG 2460  
QY 2461 AGCTATTTTAAAGCTTAGAGTTGAAGCAAAATTTGACCAAGCAATATAGTGTATTTA 2520  
Db 2461 AGCTATTTTAAAGCTTAGAGTTGAAGCAAAATTTGACCAAGCAATATAGTGTATTTA 2520  
QY 2521 CTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTATGAAATGGAAACAGAGCT 2580  
Db 2521 CTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTATGAAATGGAAACAGAGCT 2580  
QY 2581 TACTTATTTCTTGAACCAAGGCTTCTCTAGTGAAGCAATGAGCCTTTTGGAAACAC 2640  
Db 2581 TACTTATTTCTTGAACCAAGGCTTCTCTAGTGAAGCAATGAGCCTTTTGGAAACAC 2640  
QY 2641 AGATGTATACAAATAGCAAAAGTTTCAATTTCTTTTACCTTGAACCTGAAAGTAAATCAGTGA 2700  
Db 2641 AGATGTATACAAATAGCAAAAGTTTCAATTTCTTTTACCTTGAACCTGAAAGTAAATCAGTGA 2700  
QY 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGGTTC 2760  
Db 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGGTTC 2760  
QY 2761 ATCAGCTTTTAAATAGTGTGTTTCCCATCAAACTTGTTCAGACTTATATTTCTAA 2820  
Db 2761 ATCAGCTTTTAAATAGTGTGTTTCCCATCAAACTTGTTCAGACTTATATTTCTAA 2820  
QY 2821 AACCTGGACAAAACCAATATAAACAAGTCTGCTTCTTCAAAATCTGCTG 2877  
Db 2821 AACCTGGACAAAACCAATATAAACAAGTCTGCTTCTTCAAAATCTGCTG 2877

RESULT 7

```
US-10-094-749-984
; Sequence 984, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 984
; LENGTH: 3268
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-094-749-984

Query Match          50.0%; Score 2739; DB 15; Length 3268;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3039; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2386 GTTGTGTTTCTGCCCTAGACAGTCAATCAAATTTGGGTGTCATTTCTTATGGTGTTCCTCA 2445
Db 220 GTTGTGTTTCTGCCCTAGACAGTCAATCAAATTTGGGTGTCATTTCTTATGGTGTTCCTCA 279

QY 2446 TGACGACTTTGGAAGCAGTATTTTAAAGCTAGAGTTGAAGCGAGAAATTGACCAAGCAG 2505
Db 280 TGACGACTTTGGAAGCAGTATTTTAAAGCTAGAGTTGAAGCGAGAAATTGACCAAGCAG 339

QY 2506 ATTATAGTGTATTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAATCTTTTGTATG 2565
Db 340 ATTATAGTGTATTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAATCTTTTGTATG 399

QY 2566 AAATGGAACAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGA 2625
Db 400 AAATGGAACAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGA 459

QY 2626 GCCTTTGGAACAACAGAGTATACAATAGCAAGTTTCAATCTTTTACCTTGAACGCTG 2685
Db 460 GCCTTTGGAACAACAGAGTATACAATAGCAAGTTTCAATCTTTTACCTTGAACGCTG 519

QY 2686 AAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTT 2745
Db 520 AAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTT 579

QY 2746 TTATGTTTTTGTTCATCACCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTCCAG 2805
Db 580 TTATGTTTTTGTTCATCACCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTTCAG 639

QY 2806 ACTTATATTTTCTAAACCTGAGACAAACACATATAAATAACAAACAGTCTGCTTCTTC 2865
Db 640 ACTTATATTTTCTAAACCTGAGACAAACACATATAAATAACAAACAGTCTGCTTCTTC 699

2866 AAAATTTCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGACATAA 2925
700 AAAATTTCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGACATAA 759
2926 TGGTGACGATGATTAATGACAGTGAATGATCGGTGGCTCCCATAGTGGGCTTTAA 2985
760 TGGTGACGATGATTAATGACAGTGAATGATCGGTGGCTCCCATAGTGGGCTTTAA 819
2986 ATGTGATGCAATTCAGAAAGGACTATGTTTTTGGAGCTGTTTTCAACAGTACTATGGTTT 3045
820 ATGTGATGCAATTCAGAAAGGACTATGTTTTTGGAGCTGTTTTCAACAGTACTATGGTTT 879
3046 ATTCTTTTACCTATATTAGTGAATATCATTTAGTAACTACTATCTTTTATCAATTTAAATGTGA 3105
880 ATTCTTTTACCTATATTAGTGAATATCATTTAGTAACTACTATCTTTTATCAATTTAAATGTGA 939
3106 CTGAAACCAATCCAGATCTGGAGTACCCCATCTTTTCAGAAATTTACTGATATAGTTTTTA 3165
940 CTGAAACCAATCCAGATCTGGAGTACCCCATCTTTTCAGAAATTTACTGATATAGTTTTTA 999
3166 AAATTGAGCTGTATTTTCAAGCAGCTTTGCTTTGGAAATCATTTGTTTACTGCAATGCCACCTT 3225
1000 AAATTGAGCTGTATTTTCAAGCAGCTTTGCTTTGGAAATCATTTGTTTACTGCAATGCCACCTT 1059
3226 ACTTTGCCATGGAAATGACAGAAATCATAGATCAAGCTTATATCTCAACTTAACTTT 3285
1060 ACTTTGCCATGGAAATGACAGAAATCATAGATCAAGCTTATATCTCAACTTAACTTT 1119
3286 CAGGCTTTTGGCATCTGCATATTTGATTGGACAGCTGTTGTTGATATCCCTTATTTT 3345
1120 CAGGCTTTTGGCATCTGCATATTTGATTGGACAGCTGTTGTTGATATCCCTTATTTT 1179
3346 TTATCAATCTTATTTTGTATGCTAGGAGCTTATTTGGCATTTCAATTTATGATATATTTT 3405
1180 TTATCAATCTTATTTTGTATGCTAGGAGCTTATTTGGCATTTCAATTTATGATATATTTT 1239
3406 ATACTGTAAGTTCCCTTGTGTTGTTTTTGGCTTATTGTTTATGTTTCCATCAGTTATTC 3465
1240 ATACTGTAAGTTCCCTTGTGTTGTTTTTGGCTTATTGTTTATGTTTCCATCAGTTATTC 1299
3466 TGTTCATTTATATTTGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGT 3525
1300 TGTTCATTTATATTTGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGT 1359
3526 CATTTATCTTCTGTGGCAGGTTGGCTTGTATTGCAATCACTGAAATACCTTCTTTA 3585
1360 CATTTATCTTCTGTGGCAGGTTGGCTTGTATTGCAATCACTGAAATACCTTCTTTA 1419
3586 TGGGATACACAATTCGAATCTTCTTCAATTTATGCTTTTGTATCATTCATTCATCTATC 3645
1420 TGGGATACACAATTCGAATCTTCTTCAATTTATGCTTTTGTATCATTCATTCATCTATC 1479
3646 CACTTCTAGTTGCTGATTTCTTTTCAATAAGATTTCTTGGAGAAATGTACCAAAAAATG 3705
1480 CACTTCTAGTTGCTGATTTCTTTTCAATAAGATTTCTTGGAGAAATGTACCAAAAAATG 1539
3706 TGGACACCTTATTAATCCATGGGATAGGCTTTCTAGTAGTGTATTATCGCCTTACCTGCACT 3765
1540 TGGACACCTTATTAATCCATGGGATAGGCTTTCTAGTAGTGTATTATCGCCTTACCTGCACT 1599
3766 GTGTACTGTGGATTTCTTCTTCAATATCTATGAGAAAAAATATGGAGGAGATCAATTA 3825
1600 GTGTACTGTGGATTTCTTCTTCAATATCTATGAGAAAAAATATGGAGGAGATCAATTA 1659
3826 GAAAAGATCCCTTTTTCAGAAAACCTTTTCAACGAAAGCTTAAAAATAGGAAGCTTCCAGAAC 3885
1660 GAAAAGATCCCTTTTTCAGAAAACCTTTTCAACGAAAGCTTAAAAATAGGAAGCTTCCAGAAC 1719
3886 CACCAGACAATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGAGCTAAAGGTCAAGAGC 3945
1720 CACCAGACAATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGAGCTAAAGGTCAAGAGC 1779
```

3946	Qy	7GATGGGTTGCCAGTGTGTGTGAGGAGAAACCATCCATATATGGTCAGCAATTTGCAATAAG	4005
1780	Db	7GATGGGTTGCCAGTGTGTGTGAGGAGAAACCATCCATATATGGTCAGCAATTTGCAATAAG	1839
4006	Qy	AATATGATGACAGAAAGATTTTCTCTTTCAAGAAAGTAAAGAAAGTCGCAACTAAAT	4065
1840	Db	AATATGATGACAGAAAGATTTTCTCTTTCAAGAAAGTAAAGAAAGTCGCAACTAAAT	1899
4066	Qy	ACATCTCTTCTGTGTGAAAGAGAGATCTTTPAGGACTATTTGGGTCGCAAAATGGTGCTG	4125
1900	Db	ACATCTCTTCTGTGTGAAAGAGAGATCTTTPAGGACTATTTGGGTCGCAAAATGGTGCTG	1959
4126	Qy	GCAAAAGCAAAATTAATTAATTTCTGTTCTGTCGTCGATATTGGAACCAACTTCAGGCCAGGTAT	4185
1960	Db	GCAAAAGCAAAATTAATTAATTTCTGTTCTGTCGTCGATATTGGAACCAACTTCAGGCCAGGTAT	2019
4186	Qy	TTTTAGGAGATTATTTCTTCAGAGCAAGTGAAGATGATGATTCACCTGAAGTGTATGGGTT	4245
2020	Db	TTTTAGGAGATTATTTCTTCAGAGCAAGTGAAGATGATGATTCACCTGAAGTGTATGGGTT	2079
4246	Qy	ACTGTCTCAGATAAACCCCTTTGTGGCCAGATCTACATTTGCGAGGACATTTTGAATTT	4305
2080	Db	ACTGTCTCAGATAAACCCCTTTGTGGCCAGATCTACATTTGCGAGGACATTTTGAATTT	2139
4306	Qy	ATGGAGCTGTCAAAGGAATCAGTGCAAAGTGACATGAAAGAAAGTCATAAGTCGAATAACAC	4365
2140	Db	ATGGAGCTGTCAAAGGAATCAGTGCAAAGTGACATGAAAGAAAGTCATAAGTCGAATAACAC	2199
4366	Qy	ATGCATCTGATTTAAAGAACATCTTCAGAAAGACTGTAAAGAACTACTTCGAGGAATCA	4425
2200	Db	ATGCATCTGATTTAAAGAACATCTTCAGAAAGACTGTAAAGAACTACTTCGAGGAATCA	2259
4426	Qy	AACGAAAGTTGTGTTTGGCTCTAAGTATGCTTAGGGAATCCTCGAGATPACTTTGCTAGATG	4485
2260	Db	AACGAAAGTTGTGTTTGGCTCTAAGTATGCTTAGGGAATCCTCGAGATPACTTTGCTAGATG	2319
4486	Qy	AACCATCTACAGTATGGATCCCAAGCCAAAACAGCACTGTGGCGGAGCAATTCGAACTG	4545
2320	Db	AACCATCTACAGTATGGATCCCAAGCCAAAACAGCACTGTGGCGGAGCAATTCGAACTG	2379
4546	Qy	CATTTAAAAACAGAAAGCGGCTGCTATTCTGTGACCTCACTATATCGAGGAGCAGAGG	4605
2380	Db	CATTTAAAAACAGAAAGCGGCTGCTATTCTGTGACCTCACTATATCGAGGAGCAGAGG	2439
4606	Qy	CTGTCTGTGATCGAGTAGCTATCATGTGTCTGGCAGTTTAAAGATGTATCGGAACAGTAC	4665
2440	Db	CTGTCTGTGATCGAGTAGCTATCATGTGTCTGGCAGTTTAAAGATGTATCGGAACAGTAC	2499
4666	Qy	AACATCTAAAGAGTAAATTTGGRAAGGCTACTTTTTGGAAATTTAAATTTGAAAGCTGGA	4725
2500	Db	AACATCTAAAGAGTAAATTTGGRAAGGCTACTTTTTGGAAATTTAAATTTGAAAGCTGGA	2559
4726	Qy	TAGAAAACCTAGAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAAAATGCAA	4785
2560	Db	TAGAAAACCTAGAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAAAATGCAA	2619
4786	Qy	GCCTGACGAAAGTTTTTCTTCTATTTTTGGCTATAAAATTCCTAAGGAGATGTTTCAGT	4845
2620	Db	GCCTGACGAAAGTTTTTCTTCTATTTTTGGCTATAAAATTCCTAAGGAGATGTTTCAGT	2679
4846	Qy	CCCTTTCCAAATCTTTTTTAAAGCTGGAAGCTAAACATGCTTTTGGCCATTTGCAAGAT	4905
2680	Db	CCCTTTCCAAATCTTTTTTAAAGCTGGAAGCTAAACATGCTTTTGGCCATTTGCAAGAT	2739
4906	Qy	ATAGCTTTTCTCAAGCAACATTTGAAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGG	4965
2740	Db	ATAGCTTTTCTCAGGCAACATTTGAAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGG	2799
4966	Qy	AGGAGATATAGTTGTGGACTTTTAAACAGCACCTTTTGGTGGGACGCAACACAGAAG	5025
2800	Db	AGGAGATATAGTTGTGGAACTTTTAAACAGCACCTTTTGGTGGGACGCAACACAGAAG	2859
5026	Qy	ATAGAGTAGTATTTTTGAATTTTGTATGTTCTGGTCTGCTTACTGGGACTCTTTCTTTTTC	5085

```

Db      2860  ATAGAGTAGTATTTTGAATTTGTTATGTCGCGTCTGCTTACTGGGACTTCTTTCTTTTTC 2919
Qy      5086  ACATAATTTTAACTTTGGTTTAAAAAGTTTTTTTATTGGAATGTTAACTGGAGAACCAAGA 5145
Db      2920  ACTTAATTTTAACTTTGGTTTAAAAAGTTTTTTTATTGGAATGTTAACTGGAGAACCAAGA 2979
Qy      5146  ACGCAGCTGAAATTTTCTTAAGCTCCTTAATGAAATGCTGTGGTGTGTGTTTTGCTTT 5205
Db      2980  ACGCAGTTGAAATTTTCTTAAGCTCCTTAATGAAATGCTGTGGTGTGTGTTTTGCTTT 3039
Qy      5206  TCTTTAAATAAAAACGTATGTAATAATTAAGTGAAGCTGCATGTTTGTATTGAAATATATTG 5265
Db      3040  TCITTAATTAACGATATGTAATAATTAAGTGAAGCTGCATGTTTGTATTGAAATATATTG 3099
Qy      5266  AACTATATAGTTGTATGTCATCTTTTTCACCATTCGAAACAGTGCCTCTGAAATTTGTG 5325
Db      3100  AACTATATAGTTGTATGTCATCTTTTTCACCATTCGAAACAGTGCCTCTGAAATTTGTG 3159
Qy      5326  ATTTAAAGGAATTTGTAATAGAAATAGTTTATTTTAAAGTTATCTTTAAAGTTATGCCATC 5385
Db      3160  ATTTAAAGGAATTTGTAATAGAAATAGTTTATTTTAAAGTTATCTTTAAAGTTATGCCATC 3219
Qy      5386  TTCTTAATTAAGTACGTAAATGTTCCAAATCTAAATATAAAAACTAAT 5430
Db      3220  TTCTTAATTAAGTACGTAAATGTTCCAAATCTAAATATAAAAACTAAT 3264

RESULT 8
US-10-090-458-3
; Sequence 3, Application US/10090458
; Publication No. US20020123107A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hongyun
; APPLICANT: Kilinski, Ligia
; APPLICANT: Le Bihan, Stephane
; TITLE OF INVENTION: NOVEL ABCAS TRANSPORTER AND USES THEREOF
; FILE REFERENCE: 100103.403
; CURRENT APPLICATION NUMBER: US/10/090,458
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4917
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; LOCATION: 2775..2776
; OTHER INFORMATION: n = A, T, C or G
US-10-090-458-3

Query Match          49.6%; Score 2713; DB 13; Length 4917;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2763; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      114  ATGTCCACTCCAAATAGGGAGGTAGGAGTTTGGAGACAGACCAGACACTTCTACTGAAG 173
Db      1      1  ATGTCCACTCGAAATAGGGAGGTAGGAGTTTGGAGACAGACCAGACACTTCTACTGAAG 60
Qy      174  AATTACTTAATTAATCAGAAACCAAAAGAGTAGTGTTCAGGAAATCTTTTTTCCACTA 233
Db      61  AATTACTTAATTAATCAGAAACCAAAAGAGTAGTGTTCAGGAAATCTTTTTTCCACTA 120
Qy      234  TTTTTTTTATTTGGTTAAATTAATTAATAGCATGTCATCCAAATAGAAATATGAGAA 293
Db      121  TTTTTTTTATTTGGTTAAATTAATTAATAGCATGTCATCCAAATAGAAATATGAGAA 180
Qy      294  GTGCTTAATATAGAACTCAATCTCTATGGCAAGTTTACTCTTTCTAATCTAATCTTTGGA 353
Db      181  GTGCTTAATATAGAACTCAATCTCTATGGCAAGTTTACTCTTTCTAATCTAATCTTTGGA 240
Qy      354  TATACTCCAGTGACTAATATTATACAGCAGCATCATCGAGAAAGTGCTACTGATCATCTA 413

```

Db	241	TATACCTCCAGTGTACTAATAATTACAGCAGCATCATGCAGAAAGTGCTACTGTACATCTA	300
Qy	414	CCTGATGTCTATAATTACTGAAGAATAACAAATGAAAGAAAGAAATGTTAAACATCCAGTCTC	473
Db	301	CCTGATGTCTATAATTACTGAAGAATAACAAATGAAAGAAAGAAATGTTAAACATCCAGTCTC	360
Qy	474	TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAGACCTCCATGTGCTTATGAACCTTCGT	533
Db	361	TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAGACCTCCATGTGCTTATGAACCTTCGT	420
Qy	534	TTTTTTTCCCTGATATGATCCAGTATCTTTATTTATATGGATTAACAGAGCTGGCTGTCTCA	593
Db	421	TTTTTTTCCCTGATATGATTCAGTATCTTTCTATTTATATGGATTCAGAGCTGGCTGTCTCA	480
Qy	594	AAATCATGTGAGCGTCTCAGTACTGGTCTCAGGTTTTTCACAGTTTACAAGCATCCATA	653
Db	481	AAATCATGTGAGCGTCTCAGTACTGGTCTCAGGTTTTTCACAGTTTACAAGCATCCATA	540
Qy	654	GATGCTGCCAATTATACAGTTTGAAGACCAATGTTTCTTTTGGAGAGAGCTGGAGTCAACT	713
Db	541	GATGCTGCCAATTATACAGTTTGAAGACCAATGTTTCTTTTGGAGAGAGCTGGAGTCAACT	600
Qy	714	AAAGCTGTTTATTATGGAGAGAACTGCTGTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA	773
Db	601	AAAGCTGTTTATTATGGAGAGAACTGCTGTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA	660
Qy	774	ATTTTAATATACCTAGTTATAGCAATTTTACCTTTTGGATACATTTTGGCAATTCATATC	833
Db	661	ATTTTAATATACCTAGTTATAGCAATTTTACCTTTTGGATACATTTTGGCAATTCATATC	720
Qy	834	GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAATAGGACATTCATGATAC	893
Db	721	GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAATAGGACATTCATGATAC	780
Qy	894	GCCTTTTGGCTTTCCCTGGTCTTCTATATACAGTTTAAATTTTCTTATGTGCCCTTCTT	953
Db	781	GCCTTTTGGCTTTCCCTGGTCTTCTATATACAGTTTAAATTTTCTTATGTGCCCTTCTT	840
Qy	954	ATGCGAGTCAATGGCAGAGCTCTCTTCTGTTATTTTCCCTCAAAGTAGCAGCATGTGATATTT	1013
Db	841	ATGCGAGTCAATGGCAGAGCTCTCTTCTGTTATTTTCCCTCAAAGTAGCAGCATGTGATATTT	900
Qy	1014	CTGCTTTTTTCCCTTTATGGATTATCATCTGTATTTTTTCCCTTAATGTGCACCTCTTT	1073
Db	901	CTGCTTTTTTCCCTTTATGGATTATCATCTGTATTTTTTCCCTTAATGTGCACCTCTTT	960
Qy	1074	TTTTAAAAATCAAAACATGTGGGAATAGTTGAAATTTTCTGTTACTGTGGCTTTTGGATTT	1133
Db	961	TTTTAAAAATCAAAACATGTGGGAATAGTTGAAATTTTCTGTTACTGTGGCTTTTGGATTT	1020
Qy	1134	ATTGGCCCTTATGATAATCCTCATAGAAAGTTTTCCCAAATCGTTAGTGTGGCTTTTCAGT	1193
Db	1021	ATTGGCCCTTATGATAATCCTCATAGAAAGTTTTCCCAAATCGTTAGTGTGGCTTTTCAGT	1080
Qy	1194	GCCTTCTGTCACTGTACTTTTGTGATTTGTTATTTGCACAGGTCATGCAATTTTAGAGATTTT	1253
Db	1081	GCCTTCTGTCACTGTACTTTTGTGATTTGTTATTTGCACAGGTCATGCAATTTTAGAGATTTT	1140
Qy	1254	AATGAAGGTGCTTCATTTTTCAAAATTTGACATGCGAGGCCCATATCCTCTAAATTAATTACAATT	1313
Db	1141	AATGAAGGTGCTTCATTTTTCAAAATTTGACATGCGAGGCCCATATCCTCTAAATTAATTACAATT	1200
Qy	1314	ATCATGTCTCACATTAATAGTATATTTCTATGTGCTCTTGGCTGTCTATCTTGATCAAGTC	1373
Db	1201	ATCATGTCTCACATTAATAGTATATTTCTATGTGCTCTTGGCTGTCTATCTTGATCAAGTC	1260
Qy	1374	ATTTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATTTTCTGAAAGCCTTCATATGG	1433
Db	1261	ATTTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATTTTCTGAAAGCCTTCATATGG	1320
Qy	1434	TCAAGAGCAGAAAGAAATTAATGAGGAGTTATACAGAGGCAATGTTAATGGAAATPATTAGT	1493

QY	2574	CAGAGCTTACTTATTCTTTCTGAAACCAGGCTTCTCTAGTGAGCACCATTGAGCCCTTTGG	2633
Db	2461	CAGAGCTTACTTATTCTTTCTGAAACCAGGCTGCTCTAGTGAGCACCATTGAGCCCTTTGG	2520
QY	2634	AAACAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTTACCTTTGAAACGTGAAAGTAAA	2693
Db	2521	AAACAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTTACCTTTGAAACGTGAAAGTAAA	2580
QY	2694	TCAGTGAGATFCAGTGTGTTCTCTGCTTTTAAATTTTTTTTTCACAGTTCAGATTTTTATGTTTT	2753
Db	2581	TCAGTGAGATFCAGTGTGTTCTCTGCTTTTAAATTTTTTTTTCACAGTTCAGATTTTTATGTTTT	2640
QY	2754	TTGGTTTCACTCTCTTTTAAAAATGCTGTGGTTCCCATCAAACCTGCTTCAGACTTTATAT	2813
Db	2641	TTGGTTTCACTCTCTTTTAAAAATGCTGTGGTTCCCATCAAACCTGCTTCAGACTTTATAT	2700
QY	2814	TTTTCTAAAACCTGGAGACAAACCACATAAATACAAACCAAGTCGTCTTCTTCAAAATCTT	2873
Db	2701	TTTTCTAAAACCTGGAGACAAACCACATAAATACAAACCAAGTCGTCTTCTTCAAAATCTT	2760
QY	2874	GCTG 2877	
Db	2761	GCTG 2764	
RESULT 9			
US-10-108-260A-160			
; Sequence 160, Application US/10108260A			
; Publication No. US20040005560A1			
; GENERAL INFORMATION:			
; APPLICANT: HELIX RESEARCH INSTITUTE			
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA			
; FILE REFERENCE: HL-A0106			
; CURRENT APPLICATION NUMBER: US/10/108,260A			
; CURRENT FILING DATE: 2002-03-27			
; NUMBER OF SEQ ID NOS: 5458			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 160			
; LENGTH: 3347			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-108-260A-160			
Query Match 36.4%; Score 1991; DB 15; Length 3347;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 2141; Conservative 0; Mismatches 3; Indels 0; Gaps 0;			
QY	3287	AGGTCTTTTGCCATCTGCATATTGGATTGGACGAAGCTGTTGTGATATCCCTTATTTTT	3346
Db	1171	AGGTCTTTTGCCATCTGCATATTGGATTGGACGAAGCTGTTGTGATATCCCTTATTTTT	1230
QY	3347	TATCATTTCTATTGTTGATGCTAGGAAGCTTATTGGCATTTTCATTATGGATTATATTTTA	3406
Db	1231	TATCATTTCTATTGTTGATGCTAGGAAGCTTACTGGCATTTTCATTATGGATTATATTTTA	1290
QY	3407	TACTGTAAGTTCCTGCTGTGGTTTTTTTGCCCTTATTTGGTTATGTTCCATFCAGTTATCT	3466
Db	1291	TACTGTAAGTTCCTGCTGTGGTTTTTTTGCCCTTATTTGGTTATGTTCCATFCAGTTATCT	1350
QY	3467	GTTTCACATTATATGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTC	3526
Db	1351	GTTTCACATTATATGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTC	1410
QY	3527	ATTTTATCTATTCTGTGGCAGGTTGGCTTGTTATTTGCAATCACTGAAATAACTTCTTTAT	3586
Db	1411	ATTTTATCTATTCTGTGGCAGGTTGGCTTGTTATTTGCAATCACTGAAATAACTTCTTTAT	1470
QY	3587	GGGATACAAATTCGAACACTATTTCTCATATTATGCTTTTTTGTATCATCATTTCCATCTATCC	3646
Db	1471	GGGATACAAATTCGAACACTATTTCTTCATATTATGCTTTTTTGTATCATCATTTCCATCTATCC	1530
QY	3647	ACTTCTAGGTTGCTCGATTTTCTTTCATPAAAGATTTCTTTGGAAGAAATGACGAAAAAANGT	3706













Search completed: April 12, 2004, 21:29:49  
Job time : 1252 secs

Db	1802	AGGAGCTGGAGTCAACTAAAGCTGTTATATGGGAGAACTGCTGTTGTAGAAATAGATA	1861
Qy	757	CCATTCCCCGAGGAGTAGTAATTTTAAATACCTAGTTATAGCAATTTCCACCTTTTGGATACT	816
Db	1862	CCATTCCCCGAGGAGTAGTAATTTTAAATACCTAGTTATAGCAATTTCCACCTTTTGGATACT	1921
Qy	817	TTTTGGCAATTCATATCGTAGC	838
Db	1922	TTTTGGCAATTCATATCGTAGC	1943

RESULT 15  
US-10-085-783A-44227  
; Sequence 44227, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 8994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 44227  
; LENGTH: 465  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-44227

Query Match	8.4%;	Score 459;	DB 12;	Length 465;
Best Local Similarity	100.0%;	Prod. No. 8.8e-209;		
Matches 459;	Conservative 0;	Mismatches 20;	Indels 0;	Gaps 0;
4648	GATGTATCGGAACAGTACCAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAA	4707		
Db	7 GATGTATCGGAACAGTACCAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAA	66		
4708	TTAAATTGAAGGACTGATAGAAAACCTAGAACTAGAACCGCTTCGAAGAGAATTCAGT	4767		
Db	67 TTAAATTGAAGGACTGATAGAAAACCTAGAACTAGAACCGCTTCGAAGAGAATTCAGT	186		
4768	ATATTTTCCCAAAATGCAAGCGCTCAGGAAAGCTTTTCTCTATTTTGGCTTTATAAAATTC	4827		
Db	127 ATATTTTCCCAAAATGCAAGCGCTCAGGAAAGCTTTTCTCTATTTTGGCTTTATAAAATTC	186		
4828	CTAAGGAAGATGTTTCAGTCCCTTTCTCAAACTTTTTTAACTGGAGAAGCTTAAACATG	4887		
Db	187 CTAAGGAAGATGTTTCAGTCCCTTTCTCAAACTTTTTTAACTGGAGAAGCTTAAACATG	246		
4888	CTTTTGCCATTGAAGAAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAAC	4947		
Db	247 CTTTTGCCATTGAAGAAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAAC	306		
4948	TCACTAAAGAACAGAGAGGAGGAAGATATAGTTGTGGAACTTTTAAACAGACACTTTGGT	5007		
Db	307 TCACCTAAAGAACAGAGAGGAGGAAGATATAGTTGTGGAACTTTTAAACAGACACTTTGGT	366		
5008	GGGAACGAAACAAGAAGATAGAGTAGTATTTTCAATTTGTATTTGTTTCGGTCTGCTTACT	5067		
Db	367 GGGAAACGAAACAAGAAGATAGAGTAGTATTTTCAATTTGTATTTGTTTCGGTCTGCTTACT	426		
5068	GGGACTTCTTTCTTTTTCACCTAAATTTTAACTTTGGTTTT	5106		
Db	427 GGGACTTCTTTCTTTTTCACCTAAATTTTAACTTTGGTTTT	465		

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2004, 11:03:41 ; Search time 252 Seconds  
(without alignments)  
12056.969 Million cell updates/sec

Title: US-10-090-458-4  
Perfect score: 5475  
Sequence: 1 gcgtcgccctcgacag.....cataagcaatgtgaagtt 5475

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfilesi.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	743	13.6	974	4	US-09-833-381-1085
2	318	5.8	485	4	US-09-621-976-409
3	21	0.4	2198	4	US-09-620-312D-852
4	20	0.4	693	4	US-09-328-352-3244
5	20	0.4	837	4	US-09-891-641-36
6	20	0.4	970	4	US-09-016-434-1416
7	20	0.4	1100	1	US-09-196-350-2
8	20	0.4	2433	4	US-09-540-824-24
9	20	0.4	4079	4	US-09-016-434-1247
10	20	0.4	112132	4	US-09-741-150-3
11	20	0.4	112132	4	US-10-160-187-3
12	20	0.4	1664976	4	US-08-916-421B-1
13	19	0.3	210	4	US-09-016-434-426
14	19	0.3	369	4	US-09-621-976-8464
15	19	0.3	634	1	US-08-450-065-1
16	19	0.3	634	1	US-08-450-595-1
17	19	0.3	893	4	US-09-370-838-166
18	19	0.3	2017	1	US-07-667-276A-3
19	19	0.3	3001	4	US-09-539-333D-171
20	19	0.3	7653	2	US-08-394-189B-1
21	19	0.3	7653	3	US-08-258-287B-1
22	19	0.3	7653	3	US-08-368-704C-1
23	19	0.3	7653	5	PCT-US93-05701-18
24	19	0.3	7653	5	PCT-US93-05705-1
25	19	0.3	8312	4	US-09-620-312D-1048
26	19	0.3	9370	1	US-08-320-559-27
27	19	0.3	9370	3	US-08-545-860D-27

C 28	19	0.3	9370	5	PCT-US94-04496-27	Sequence 27, Appl
C 29	19	0.3	9391	1	US-08-320-559-25	Sequence 25, Appl
C 30	19	0.3	9391	3	US-08-545-860D-25	Sequence 25, Appl
C 31	19	0.3	9391	5	PCT-US94-04496-25	Sequence 25, Appl
C 32	19	0.3	46718	4	US-09-816-093-3	Sequence 3, Appl
C 33	19	0.3	148567	4	US-09-801-876B-3	Sequence 3, Appl
C 34	19	0.3	148567	4	US-10-254-869-3	Sequence 3, Appl
C 35	19	0.3	269223	4	US-09-596-002-41	Sequence 41, Appl
C 36	19	0.3	640681	4	US-09-790-988-1	Sequence 1, Appl
C 37	19	0.3	1830121	4	US-09-557-884-1	Sequence 1, Appl
C 38	19	0.3	1830121	4	US-09-643-990A-1	Sequence 1, Appl
C 39	18	0.3	59	4	US-09-532-656-2	Sequence 2, Appl
C 40	18	0.3	163	2	US-08-611-757-78	Sequence 78, Appl
C 41	18	0.3	163	5	PCT-US95-05980-78	Sequence 78, Appl
C 42	18	0.3	231	4	US-09-134-001C-1450	Sequence 1450, Ap
C 43	18	0.3	387	4	US-09-134-001C-1867	Sequence 1867, Ap
C 44	18	0.3	474	4	US-09-328-352-39	Sequence 39, Appl
C 45	18	0.3	525	4	US-09-134-001C-1623	Sequence 1623, Ap

ALIGNMENTS

RESULT 1  
US-09-833-381-1085  
; Sequence 1085, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1085  
; LENGTH: 974  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1085

Query Match	13.6%;	Score	743;	DB	4;	Length	974;
Best Local Similarity	100.0%;	Pred. No.	0;	Mismatches	0;	Indels	0;
Matches	743;	Conservative	0;				
QY	97	AGGTTTATTCGAAACATGTCCTGCAATTTAGGAGGTAGGAGTTTGGAGACAGACCA	156				
Db	216	AGGTTTATTCGAAACATGTCCTGCAATTTAGGAGGTAGGAGTTTGGAGACAGACCA	275				
QY	157	GAACACTTCTACTGAAGAAATACCTTTAATAATGCAGAACCAAAAGAGTAGTGTTCAGG	216				
Db	276	GAACACTTCTACTGAAGAAATACCTTTAATAATGCAGAACCAAAAGAGTAGTGTTCAGG	335				
QY	217	AAATCTTTTCCACTATTTTATTTTGGTAAATTAATTAATAGCATGATGCATCCAA	276				
Db	336	AAATCTTTTCCACTATTTTATTTTGGTAAATTAATTAATAGCATGATGCATCCAA	395				
QY	277	ATAAGAAATATGAAGAGTGCTTAATATAGAACTCAATCTTATGGAAGTATCTCTTT	336				
Db	396	ATAAGAAATATGAAGAGTGCTTAATATAGAACTCAATCTTATGGAAGTATCTCTTT	455				
QY	337	CTAATCTAATCTTGGATATATCTCCAGTGACTTAATATTACAGCAGCATGATGCAGAAAG	396				
Db	456	CTAATCTAATCTTGGATATATCTCCAGTGACTTAATATTACAGCAGCATGATGCAGAAAG	515				
QY	397	TGCTTACTGATCATCTACCTGATGTGCTAATAATTACTGAAGAATATACAATGAAAAAGAAA	456				
Db	516	TGCTTACTGATCATCTACCTGATGTGCTAATAATTACTGAAGAATATACAATGAAAAAGAAA	575				
QY	457	TGTTAACTCCAGTCTCTTAAGCCGACACTTTGAGGTGTGTTTTTCAAGACTCCA	516				

Db 576 TGTAAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGTGTTTTCAAGACTCCA 635  
 QY 517 TGTCTATCAACTTCGTTTTTTCTGTATATGATCCAGTATCTTCTATTATATGGATT 576  
 Db 636 TGTCTATGAACCTTGTTTTTTTCTGTATATGATCCAGTATCTTCTATTATATGGATT 695  
 QY 577 CAAGAGCTGGCTGTCTCAAAATCATGTGAGGCTGCTCAGTACTGTGTCCTCAGGTTTCACG 636  
 Db 696 CAAGAGCTGGCTGTCTCAAAATCATGTGAGGCTGCTCAGTACTGTGTCCTCAGGTTTCACG 755  
 QY 637 TTTTCAAGCATCCATAGATGCTGCATTATACAGTTCGAAGCAAACTTTCTCTTTGGA 696  
 Db 756 TTTTCAAGCATCCATAGATGCTGCATTATACAGTTCGAAGCAAACTTTCTCTTTGGA 815  
 QY 697 AGGAGCTGGAGTCAACTAAAGCTGTATTATGGGAGAAACTGCTGTTGTAGAAATAGATA 756  
 Db 816 AGGAGCTGGAGTCAACTAAAGCTGTATTATGGGAGAAACTGCTGTTGTAGAAATAGATA 875  
 QY 757 CTTTCCCGGAGGAGTAAATTAATATACCTAGTTATAGCAATTTTCACCTTTTGATACT 816  
 Db 876 CTTTCCCGGAGGAGTAAATTAATATACCTAGTTATAGCAATTTTCACCTTTTGATACT 935  
 QY 817 TTTTGGCAATTCATCTGTAGCA 839  
 Db 936 TTTTGGCAATTCATCTGTAGCA 958

RESULT 2

US-09-621-976-409  
 ; Sequence 409, Application US/09621976  
 ; Patent No. 6639063  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 409  
 ; LENGTH: 485  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; NAME/KEY: CDS  
 ; LOCATION: 185..484  
 US-09-621-976-409

Query Match 5.8%; Score 318; DB 4; Length 485;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-140;  
 Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 AGTATTATTCAGAAACATGTCCTCAATTTAGGAGGTAGGAGTTGGAGACAGACCA 156  
 Db 168 AGTATTATTCAGAAACATGTCCTCAATTTAGGAGGTAGGAGTTGGAGACAGACCA 227  
 QY 157 GAACACTTCTAGTGAAGATTACTTAATTAATGCAACCAAAAGAGTAGTGTTCAGG 216  
 Db 228 GAACACTTCTAGTGAAGATTACTTAATTAATGCAACCAAAAGAGTAGTGTTCAGG 287  
 QY 217 AAATCTTTTTCAGTATTTTCTTTTATTTGTTTAAATTAATTAATGATGATGATCAAA 276  
 Db 288 AAATCTTTTTCAGTATTTTCTTTTATTTGTTTAAATTAATTAATGATGATGATCAAA 347  
 QY 277 ATAAGAAATATGAAGAGTCCCTAATATAGAACTCAATCCCTATGACAAAGTTTACTCTTT 336  
 Db 348 ATAAGAAATATGAAGAGTCCCTAATATAGAACTCAATCCCTATGACAAAGTTTACTCTTT 407  
 QY 337 CTAATCTAATTTCTTGGATATCTCCAGTGCCTAATTAATTAATGAGGAGTATGAGGAAAG 396  
 Db 408 CTAATCTAATTTCTTGGATATCTCCAGTGCCTAATTAATTAATGAGGAGTATGAGGAAAG 467

QY 397 TGTCTACTGATCATCTAC 414  
 Db 468 TGTCTACTGATCATCTAC 485

RESULT 3

US-09-620-312D-852  
 ; Sequence 852, Application US/09620312D  
 ; Patent No. 6569662  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Liu, Chenghua  
 ; APPLICANT: Asundi, Vinod  
 ; APPLICANT: Zhang, Jie  
 ; APPLICANT: Ren, Feiyan  
 ; APPLICANT: Chen, Rui-hong  
 ; APPLICANT: Zhao, Qing A.  
 ; APPLICANT: Wehrman, Tom  
 ; APPLICANT: Xue, Aidong J.  
 ; APPLICANT: Yang, Yonghong  
 ; APPLICANT: Wang, Jian-Rui  
 ; APPLICANT: Zhou, Ping  
 ; APPLICANT: Ma, Yunging  
 ; APPLICANT: Wang, Dunrui  
 ; APPLICANT: Wang, Zhiwei  
 ; APPLICANT: John Tillinghast  
 ; APPLICANT: Drmanac, Radoje T.  
 ; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
 ; FILE REFERENCE: 784CIP2B  
 ; CURRENT APPLICATION NUMBER: US/09/620,312D  
 ; CURRENT FILING DATE: 2000-07-19  
 ; PRIOR APPLICATION NUMBER: 09/552,317  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 09/488,725  
 ; PRIOR FILING DATE: 2000-01-21  
 ; NUMBER OF SEQ ID NOS: 1105  
 ; SOFTWARE: pt FL\_genes Version 1.0  
 ; SEQ ID NO 852  
 ; LENGTH: 2198  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (206)..(1867)  
 US-09-620-312D-852

Query Match 0.4%; Score 21; DB 4; Length 2198;  
 Best Local Similarity 100.0%; Pred. No. 7.5;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3895 ATGAGGATGAAGATGAAGATG 3915  
 Db 1503 ATGAGGATGAAGATGAAGATG 1523

RESULT 4

US-09-328-352-3244  
 ; Sequence 3244, Application US/09328352  
 ; Patent No. 6562958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328,352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 3244  
 ; LENGTH: 693  
 ; TYPE: DNA  
 ; ORGANISM: Acinetobacter baumannii

```
US-09-328-352-3244
Query Match          0.4%; Score 20; DB 4; Length 693;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3901 ATGAGATGAAGATCTCAA 3920
Db 497 ATGAAGATGAAGATCTCAA 516

RESULT 5
US-09-891-641-36/c
; Sequence 36, Application US/09891641
; Patent No. 6617148
; GENERAL INFORMATION:
; APPLICANT: Ye, Rick
; APPLICANT: Bedzyk, Laura
; APPLICANT: Wang, Tao
; TITLE OF INVENTION: NATURAL PROMOTERS FOR GENE EXPRESSION IN BACILLUS SPECIES
; FILE REFERENCE: C11686 US NA
; CURRENT APPLICATION NUMBER: US/09/891,641
; CURRENT FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft 97
; SEQ ID NO 36
; LENGTH: 837
; TYPE: DNA
; ORGANISM: Bacillus subtilis
US-09-891-641-36

Query Match          0.4%; Score 20; DB 4; Length 837;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2752 TTTGGTTTCATCACTCTTT 2771
Db 812 TTTGGTTTCATCACTCTTT 793

RESULT 6
US-09-016-434-1416/c
; Sequence 1416, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071

US-09-328-352-3244
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1416:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 970 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9531102
US-09-016-434-1416

Query Match          0.4%; Score 20; DB 4; Length 970;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1747 TCTCAGAAATAGATGAAATG 1766
Db 854 TCTCAGAAATAGATGAAATG 835

RESULT 7
US-08-196-350-2/c
; Sequence 2, Application US/08196350
; Patent No. 5585099
; GENERAL INFORMATION:
; APPLICANT: Richards, Sue
; APPLICANT: Kaplan, Joanne
; APPLICANT: Mosciaki, Richard
; TITLE OF INVENTION: PROLACTIN AS ADJUVANT
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brad Salcedo
; STREET: One Kendall Square
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02139
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/196,350
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Goez, William G
; REGISTRATION NUMBER: 27,787
; REFERENCE/DOCKET NUMBER: GEN 4-1.0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 6172527868
; TELEFAX: 6173747225
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1100 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-196-350-2

Query Match          0.4%; Score 20; DB 1; Length 1100;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1747 TCTCAGAAATAGATGAAATG 1766
Db 984 TCTCAGAAATAGATGAAATG 965
```

```

; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g190037
; US-09-016-434-1247

Query Match          0.4%; Score 20; DB 4; Length 4079;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2778 GCTGTGTTCCCATCAAACT 2797
Db 2590 GCTGTGTTCCCATCAAACT 2609

RESULT 10
US-09-741-150-3
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL0000968
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(112132)
; OTHER INFORMATION: n = A,T,C or G
; US-09-741-150-3

Query Match          0.4%; Score 20; DB 4; Length 112132;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 788 AGTTATAGCATTTTCCACTT 807
Db 55116 AGTTATAGCATTTTCCACTT 55135

RESULT 11
US-10-160-187-3
; Sequence 3, Application US/10160187
; Patent No. 6620607
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000968DIV
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/252,410
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 09/741,150
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(112132)
; OTHER INFORMATION: n = A,T,C or G
; US-10-160-187-3

; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g190037
; US-09-016-434-1247

Query Match          0.4%; Score 20; DB 4; Length 4079;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2778 GCTGTGTTCCCATCAAACT 2797
Db 2590 GCTGTGTTCCCATCAAACT 2609

RESULT 10
US-09-741-150-3
; Sequence 3, Application US/09741150
; Patent No. 6436689
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL0000968
; CURRENT FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(112132)
; OTHER INFORMATION: n = A,T,C or G
; US-09-741-150-3

Query Match          0.4%; Score 20; DB 4; Length 112132;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 788 AGTTATAGCATTTTCCACTT 807
Db 55116 AGTTATAGCATTTTCCACTT 55135

RESULT 11
US-10-160-187-3
; Sequence 3, Application US/10160187
; Patent No. 6620607
; GENERAL INFORMATION:
; APPLICANT: GUEGLER, Karl et al.
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000968DIV
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/252,410
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 09/741,150
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 112132
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(112132)
; OTHER INFORMATION: n = A,T,C or G
; US-10-160-187-3
```



US-10-160-187-3

Query Match 0.4%; Score 20; DB 4; Length 112132;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 788 AGTTATAGCATTTTCACCTT 807  
Db 55116 AGTTATAGCATTTTCACCTT 55135

RESULT 12

US-08-916-421B-1  
; Sequence 1, Application US/08916421B  
; Patent No.: 6503723  
; GENERAL INFORMATION:  
; APPLICANT: Bult et al.  
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus  
; Patent No. 6503723  
; TITLE OF INVENTION: jannaschii  
; FILE REFERENCE: PB275  
; CURRENT APPLICATION NUMBER: US/08/916,421B  
; CURRENT FILING DATE: 1997-08-22  
; PRIOR APPLICATION NUMBER: US 60/024,428  
; PRIOR FILING DATE: 1996-08-22  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1664976  
; TYPE: DNA  
; ORGANISM: Methanococcus jannaschii  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (28222)..(28222)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (28257)..(28258)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84773)..(84773)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84808)..(84808)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (84812)..(84812)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98120)..(98120)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98159)..(98159)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98239)..(98239)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98266)..(98266)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (98343)..(98343)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (103998)..(103998)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (148948)..(148948)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (163385)..(163385)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (191989)..(191989)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (191995)..(191995)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (231580)..(231580)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (234187)..(234187)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (234220)..(234220)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (234814)..(234814)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (309398)..(309398)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (309418)..(309418)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (312837)..(312837)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (312993)..(312993)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (319226)..(319226)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (559167)..(559167)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (559241)..(559241)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (600992)..(600992)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (622708)..(622708)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (657081)..(657081)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (657203)..(657203)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (674435)..(674435)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (682442)..(682442)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (713652)..(713652)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (741684)..(741684)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (779455)..(779455)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (779676)..(779676)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (855539)..(855539)  
; OTHER INFORMATION: n equals a, t, c, or g  
; NAME/KEY: misc feature  
; LOCATION: (871619)..(871619)  
; OTHER INFORMATION: n equals a, t, c, or g

NAME/KEY: misc feature  
LOCATION: (1084830)..(1084830)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1096846)..(1096846)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1119881)..(1119881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1130881)..(1130881)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1310988)..(1310988)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1313224)..(1313224)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349473)..(1349473)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1349491)..(1349491)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1470091)..(1470091)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1569020)..(1569020)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1602912)..(1602912)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1603734)..(1603734)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1637998)..(1637998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (1664854)..(1664854)  
OTHER INFORMATION: n equals a, t, c, or g  
US-08-916-421B-1

Query Match 0.4%; Score 20; DB 4; Length 1664976;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4938 TTGTAGACTCACTAAGA 4957  
DB 201642 TTGTAGACTCACTAAGA 201661

RESULT 13  
US-09-016-434-426/c  
Sequence 426, Application US/09016434  
Patent No. 6500938  
GENERAL INFORMATION:  
APPLICANT: Janice Au-Young  
APPLICANT: Jeffrey J. Seilhamer  
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING  
TITLE OF INVENTION: PATHWAY GENE EXPRESSION  
NUMBER OF SEQUENCES: 1490  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.  
STREET: 3174 PORTER DRIVE  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/016,434  
FILING DATE: HERewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0002 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 426:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 210 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: STOMNOT01  
CLONE: 215814  
US-09-016-434-426

Query Match 0.3%; Score 19; DB 4; Length 210;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1821 GTTTGACAGTAGAGAAA 1839  
DB 20 GTTTGACAGTAGAGAAA 2

RESULT 14  
US-09-621-976-8464/c  
Sequence 8464, Application US/09621976  
Patent No. 6639063  
GENERAL INFORMATION:  
APPLICANT: Dumas Milne Edwards, J.B.  
APPLICANT: Jobert, S.  
APPLICANT: Giordano, J.Y.  
TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 8464  
LENGTH: 369  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-8464

Query Match 0.3%; Score 19; DB 4; Length 369;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2709 TTGCTTCTGCTTTTAAATTT 2727  
DB 211 TTGCTTCTGCTTTTAAATTT 193

RESULT 15  
US-08-450-065-1/c  
Sequence 1, Application US/08450065  
Patent No. 5798105  
GENERAL INFORMATION:  
APPLICANT: Schoenmakers, Johannes G  
APPLICANT: Konings, Rudolph NH

APPLICANT: Moelans, Inge IMD  
TITLE OF INVENTION: No. 5798105el protein  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SmithKline Beecham Corporate Patents -US  
STREET: UW2220, Po Box 1539  
CITY: King of Prussia  
STATE: PA  
COUNTRY: USA  
ZIP: 19406-0939  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/450,065  
FILING DATE: 25-MAY-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/949645  
FILING DATE: 04-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Jervis, Herbert H.  
REGISTRATION NUMBER: 31,171  
REFERENCE/DOCKET NUMBER: B2992  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-270-5065  
TELEFAX: 215-270-5090  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 634 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: PLASMODIUM  
STRAIN: FALCIPARUM  
IMMEDIATE SOURCE:  
CLONE: 16K  
US-08-450-065-1

Query Match 0.3%; Score 19; DB 1; Length 634;  
Best Local Similarity 100.0%; Pred.No. 63;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 842 AAAAGAAAAAATAAAAA 860  
|||||  
DB 41 AAAAGAAAAAATAAAAA 23  
|||||

Search completed: April 12, 2004, 18:02:54  
Job time : 259 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: April 12, 2004, 07:56:11 ; Search time 13649 Seconds  
(without alignments)  
17386.117 Million cell updates/sec

Title: US-10-090-458-4  
Perfect score: 5475  
Sequence: 1 ggcgtccggccctcgacag.....cataaagcaatggaagtt 5475

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_ats.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_on.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_scs.\*

28: em\_un.\*

29: em\_vi.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_mam.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htg\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	5475	100.0	5475	6	AX537473	AX537473 Sequence
2	5181	94.6	6525	6	AX473847	AX473847 Sequence
3	5181	94.6	6525	9	AY028897	AY028897 Homo sapi
4	5086	92.9	6369	6	AX392931	AX392931 Sequence
5	4935	90.1	5096	9	HS4275973	AJ275973 Homo sapi
6	4664	85.2	5262	6	AX417828	AX417828 Sequence
7	4572	83.5	4529	6	AX417824	AX417824 Sequence
8	4408	80.5	4785	6	AX417826	AX417826 Sequence
9	2826	51.6	5463	6	AX537470	AX537470 Sequence
10	2739	50.0	3268	6	AX714300	AX714300 Sequence
11	2739	50.0	3268	9	AK056533	AK056533 Homo sapi
12	2713	49.6	4917	6	AX537472	AX537472 Sequence
13	2617	47.8	2845	9	HS8512612	AJ512612 Homo sapi
14	2597	47.4	3112	9	AB067475	AB067475 Homo sapi
15	2294	41.9	5243	9	HS807592	BX647447 Homo sapi
16	1991	36.4	3347	6	AX833036	AX833036 Sequence
17	1991	36.4	3347	9	AK094416	AK094416 Homo sapi
18	1758	32.1	1964	9	AK122803	AK122803 Homo sapi
19	1601	29.2	4078	9	HS806823	BX640746 Homo sapi
20	1560	28.5	1560	9	AK058170	AK058170 Homo sapi
21	1238	22.6	2325	6	AX780218	AX780218 Sequence
22	744	13.6	1818	6	AX576092	AX576092 Sequence
23	742	13.6	1943	6	AX834300	AX834300 Sequence
24	742	13.6	1943	9	AK096654	AK096654 Homo sapi
25	692	12.6	3324	9	HS808688	BX648540 Homo sapi
26	639	11.7	735	9	BC054480	BC054480 Homo sapi
27	548	10.0	548	6	AX778667	AX778667 Sequence
28	543	9.9	163229	9	AC115985	AC115985 Homo sapi
29	543	9.9	172223	2	AC115985	AC115985 Homo sapi
30	543	9.9	185252	9	AC005495	AC005495 Homo sapi
31	440	8.0	612	6	AX740455	AX740455 Sequence
32	376	6.9	1346	6	AX331591	AX331591 Sequence
33	376	6.9	1346	6	AX332287	AX332287 Sequence
34	376	6.9	1346	6	AX410760	AX410760 Sequence
35	324	5.9	469	6	AX409236	AX409236 Sequence
36	318	5.8	485	6	AR412772	AR412772 Sequence
37	318	5.8	485	6	BD108325	BD108325 EST and e
38	242	4.4	341	6	BD071308	BD071308 Secrated
39	233	4.3	71794	2	AC023647	AC023647 Homo sapi
40	233	4.3	160254	2	AC007763	AC007763 Homo sapi
41	233	4.3	178097	2	AC016704	AC016704 Homo sapi
42	227	4.1	252	6	BD071579	BD071579 Secrated
43	207	3.8	160254	2	AC007763	AC007763 Homo sapi
44	179	3.3	178097	2	AC016704	AC016704 Homo sapi
45	129	2.4	71794	2	AC023647	AC023647 Homo sapi

# ALIGNMENTS

RESULT 1  
AX537473  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

AX537473  
Sequence 4 from Patent WO2070690.  
AX537473  
AX537473.1 GI:25269282  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Chen, H., Kilinski, L. and le Bihan, S.  
Abcds transporter and uses thereof  
Patent: WO 02070690-A 4 12-SEP-2002;

AX537473  
Sequence 4 from Patent WO2070690.  
AX537473  
AX537473.1 GI:25269282  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Chen, H., Kilinski, L. and le Bihan, S.  
Abcds transporter and uses thereof  
Patent: WO 02070690-A 4 12-SEP-2002;

Pred. No. is the number of results predicted by chance to have a

\_\_\_\_\_

[REDACTED]

5

901	GGCTTTCCTGGGTCCTCTATATACAGTTTAAATTTTCTTATGTCCTCTTATGGCAG	960
961	TCATTCGGACAGCTTCCTTGTAATTTCCCTCAAGTAGACAGCATCTGTGATATTTCTGCTTT	1020
961	TCATTCGGACAGCTTCCTTGTAATTTCCCTCAAGTAGACAGCATCTGTGATATTTCTGCTTT	1020
1021	TTTTCTTTATGGATTATCATCTGTATTTTTTGTCTTTAATGCTGCACACTCTTTTAAAA	1080
1021	TTTTCTTTATGGATTATCATCTGTATTTTTTGTCTTTAATGCTGCACACTCTTTTAAAA	1080
1081	AATCAAAACATGTGGGAATAGTTGAAATTTTTTGTGTACTGTGGCTTTTGGATTTATTTGGCC	1140
1081	AATCAAAACATGTGGGAATAGTTGAAATTTTTTGTGTACTGTGGCTTTTGGATTTATTTGGCC	1140
1141	TTATGATTAATCTCTCATAGAAAGTTTCCCAAATCGTTAGTGTGGCTTTTCAGTCTTTTCT	1200
1141	TTATGATTAATCTCTCATAGAAAGTTTCCCAAATCGTTAGTGTGGCTTTTCAGTCTTTTCT	1200
1201	GTCACTGTACTTTTGTGATTTGTAATGCAAGCTCATGCAATTAGAAGATTTAATCAAG	1260
1201	GTCACTGTACTTTTGTGATTTGTAATGCAAGCTCATGCAATTAGAAGATTTAATCAAG	1260
1261	GTGCTTCATTTTCAAAATTTGACCTGACAGGCCATATCTCTTAATTTATCAATATATCAATG	1320
1261	GTGCTTCATTTTCAAAATTTGACCTGACAGGCCATATCTCTTAATTTATCAATATATCAATG	1320
1321	TCACACTTAATAGTATATTTCTATGTCTCTTGCTGTCTATCTTGTATCAAGTCAATCCAG	1380
1321	TCACACTTAATAGTATATTTCTATGTCTCTTGCTGTCTATCTTGTATCAAGTCAATCCAG	1380
1381	GGAAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCTTCATATTTGGTCAAAGA	1440
1381	GGAAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCTTCATATTTGGTCAAAGA	1440
1441	GCAAAAGAAATTTATGAGGATTTATCAGAGGCCAATGTTAATGGAATATTTAGTTTGTAGT	1500
1441	GCAAAAGAAATTTATGAGGATTTATCAGAGGCCAATGTTAATGGAATATTTAGTTTGTAGT	1500
1501	AAATATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATGAAGATTTAGTGGTA	1560
1501	AAATATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATGAAGATTTAGTGGTA	1560
1561	TTTCAGAGACATACAGAAAGAGGGTGAAATGTCGGAGCTTCGAGAAATTTCTCATTTTG	1620
1561	TTTCAGAGACATACAGAAAGAGGGTGAAATGTCGGAGCTTCGAGAAATTTCTCATTTTG	1620
1621	ACATATATGAGGGTCAGATTATCTGCCTTACTTTGGCCACAGTGGAAACAGGAAAGTACAT	1680
1621	ACATATATGAGGGTCAGATTATCTGCCTTACTTTGGCCACAGTGGAAACAGGAAAGTACAT	1680
1681	TGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCACTATATATGGAC	1740
1681	TGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCACTATATATGGAC	1740
1741	ACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAATGATTTGGCAATTTGTCCAC	1800
1741	ACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAATGATTTGGCAATTTGTCCAC	1800
1801	AGTTAGATATACACTTTTGATGTTTTGCACAGTAGAAGAAAAATTTATCAATTTTGGCTTCAA	1860
1801	AGTTAGATATACACTTTTGATGTTTTGCACAGTAGAAGAAAAATTTATCAATTTTGGCTTCAA	1860
1861	TCAAAAGGATACACAGCCAAATATATATACAGAGTGCACAGGTTTTTACTAGATTTAG	1920
1861	TCAAAAGGATACACAGCCAAATATATATACAGAGTGCACAGGTTTTTACTAGATTTAG	1920
1921	ACATCGACATATCAAAAGATATCAAGCTTAAAAATTTAAGTGTGGTGTCAAAAAGAAAGC	1980
1921	ACATCGACATATCAAAAGATATCAAGCTTAAAAATTTAAGTGTGGTGTCAAAAAGAAAGC	1980
1981	TGTCTATTAGGAATTTGCTTTCTTTGGGAACCCAAAGATACTCTCTCTAGATGAACCAACAG	2040

901	GGCTTTCCTGGGTCCTCTATATACAGTTTAAATTTTCTTATGTCCTCTTATGGCAG	960
961	TCATTCGGACAGCTTCCTTGTAATTTCCCTCAAGTAGACAGCATCTGTGATATTTCTGCTTT	1020
961	TCATTCGGACAGCTTCCTTGTAATTTCCCTCAAGTAGACAGCATCTGTGATATTTCTGCTTT	1020
1021	TTTTCTTTATGGATTATCATCTGTATTTTTTGTCTTTAATGCTGCACACTCTTTTAAAA	1080
1021	TTTTCTTTATGGATTATCATCTGTATTTTTTGTCTTTAATGCTGCACACTCTTTTAAAA	1080
1081	AATCAAAACATGTGGGAATAGTTGAAATTTTTTGTGTACTGTGGCTTTTGGATTTATTTGGCC	1140
1081	AATCAAAACATGTGGGAATAGTTGAAATTTTTTGTGTACTGTGGCTTTTGGATTTATTTGGCC	1140
1141	TTATGATTAATCTCTCATAGAAAGTTTCCCAAATCGTTAGTGTGGCTTTTCAGTCTTTTCT	1200
1141	TTATGATTAATCTCTCATAGAAAGTTTCCCAAATCGTTAGTGTGGCTTTTCAGTCTTTTCT	1200
1201	GTCACTGTACTTTTGTGATTTGTAATGCAAGCTCATGCAATTAGAAGATTTAATCAAG	1260
1201	GTCACTGTACTTTTGTGATTTGTAATGCAAGCTCATGCAATTAGAAGATTTAATCAAG	1260
1261	GTGCTTCATTTTCAAAATTTGACCTGACAGGCCATATCTCTTAATTTATCAATATATCAATG	1320
1261	GTGCTTCATTTTCAAAATTTGACCTGACAGGCCATATCTCTTAATTTATCAATATATCAATG	1320
1321	TCACACTTAATAGTATATCTATGTCTCTTGGCTGTCTATCTTGTATCAAGTCAATCCAG	1380
1321	TCACACTTAATAGTATATCTATGTCTCTTGGCTGTCTATCTTGTATCAAGTCAATCCAG	1380
1381	GGAAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCTTCATATTTGGTCAAAGA	1440
1381	GGAAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCTTCATATTTGGTCAAAGA	1440
1441	GCAAAAGAAATATGACAGGTATACAGAGGCCAATGTTAATGGAATATTTAGTTTGTAGTG	1500
1441	GCAAAAGAAATATGACAGGTATACAGAGGCCAATGTTAATGGAATATTTAGTTTGTAGTG	1500
1501	AAATATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATGAAGATTTAGTGGTA	1560
1501	AAATATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATGAAGATTTAGTGGTA	1560
1561	TTCAGAGACATACAGAAAGAGGTGAAATGTCGGAGCTTTCGAGAAATTTCTCATTTTG	1620
1561	TTCAGAGACATACAGAAAGAGGTGAAATGTCGGAGCTTTCGAGAAATTTCTCATTTTG	1620
1621	ACATATATGAGGGTCAGATTATCTGCCTTACTTTGGCCACAGTGGAAACAGGATACAT	1680
1621	ACATATATGAGGGTCAGATTATCTGCCTTACTTTGGCCACAGTGGAAACAGGATACAT	1680
1681	TGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATATATGGAC	1740
1681	TGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATATATGGAC	1740
1741	ACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAATGATTTGGCATTTGTCCAC	1800
1741	ACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAATGATTTGGCATTTGTCCAC	1800
1801	AGTTAGATATACATTTGTATGTTTTGCACAGTAGAAGAAAAATTTATCAATTTTGGCTTCAA	1860
1801	AGTTAGATATACATTTGTATGTTTTGCACAGTAGAAGAAAAATTTATCAATTTTGGCTTCAA	1860
1861	TCAAAAGGATACACAGCCAAATATATATACAGAGTGCACAGGTTTTTACTAGATTTAG	1920
1861	TCAAAAGGATACACAGCCAAATATATATACAGAGTGCACAGGTTTTTACTAGATTTAG	1920
1921	ACATCGACATATCAAAAGATACCAAGCTTAAAAATTTAAGTGTGGTTCAAAAAGAAAGC	1980
1921	ACATCGACATATCAAAAGATACCAAGCTTAAAAATTTAAGTGTGGTTCAAAAAGAAAGC	1980
1981	TGTCATTAGGAATTTGCTTTCTTTGGGAACCCAAAGATACTCTCTCTAGATGAACCAACAG	2040

1981 TGTCATTAGGAATGCTGTTCTTTGGGAACCAAGACTACTGCTGTAGATGAACCAACAG 2040  
 2041 CTGGATGGACCCCTGTTCTCGACATATGCTATGGAATCTTTTAAATACAGAAAGCCA 2100  
 2041 CTGGATGGACCCCTGTTCTCGACATATGCTATGGAATCTTTTAAATACAGAAAGCCA 2100  
 2101 ATCGGGTACAGTGTTCAGTACTCAITTCATGATGAAGCTGACATCTTTGTCAGATPAGA 2160  
 2101 ATCGGGTACAGTGTTCAGTACTCAITTCATGATGAAGCTGACATCTTTGTCAGATPAGA 2160  
 2161 AAGCTGTGATATCACAGGAATGCTGAAATGCTGTTCTTCAATGTTCTTCAAAAGTA 2220  
 2161 AAGCTGTGATATCACAGGAATGCTGAAATGCTGTTCTTCAATGTTCTTCAAAAGTA 2220  
 2221 AATGGGGGATCGGCTACCGCTCAGCATGTACATAGACAAATATTTGGCCACAGATCTC 2280  
 2221 AATGGGGGATCGGCTACCGCTCAGCATGTACATAGACAAATATTTGGCCACAGATCTC 2280  
 2281 TTTCTTCATCTGTTTAAACAATATACCTGGAGCTACTTTATACAAAGATGACCAAC 2340  
 2281 TTTCTTCATCTGTTTAAACAATATACCTGGAGCTACTTTATACAAAGATGACCAAC 2340  
 2341 AACTTGTGATAGCTTCCCTTTCAAGGACATGACAAATTTTCAGGTTTGTCTTCTGCC 2400  
 2341 AACTTGTGATAGCTTCCCTTTCAAGGACATGACAAATTTTCAGGTTTGTCTTCTGCC 2400  
 2401 TAGACATCATCAAAATTTGGGTGTCATTTCTTATGTTTCCATGACGACTTTGGAAG 2460  
 2401 TAGACATCATCAAAATTTGGGTGTCATTTCTTATGTTTCCATGACGACTTTGGAAG 2460  
 2461 ACGTATTTTAAAGCTAGAGTTGAGCAGAAATGACCAAGCAGATTTATAGTATTTA 2520  
 2461 ACGTATTTTAAAGCTAGAGTTGAGCAGAAATGACCAAGCAGATTTATAGTATTTA 2520  
 2521 CTCAGCGCCACTGGAGGAAGAAATGGATTCAAATCTTTTGTGAAATGGACAGAGCT 2580  
 2521 CTCAGCGCCACTGGAGGAAGAAATGGATTCAAATCTTTTGTGAAATGGACAGAGCT 2580  
 2581 TACTATTTCTTTGAAACCAAGCTTCTAGTGGACCAATGAGCCTTTGGAACACAC 2640  
 2581 TACTATTTCTTTGAAACCAAGCTTCTAGTGGACCAATGAGCCTTTGGAACACAC 2640  
 2641 AGATGTATACAAATAGCAAGTTTCATTTCTTTACCTTGAACCGTGAAGTAAATCAGTGA 2700  
 2641 AGATGTATACAAATAGCAAGTTTCATTTCTTTACCTTGAACCGTGAAGTAAATCAGTGA 2700  
 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAATTTTATGTTTGTGTTTC 2760  
 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAATTTTATGTTTGTGTTTC 2760  
 2761 ATCACTCTTTTAAATATGCTGTGTTTCCCATCAAACTTGTTCAGACTTATATTTCTAA 2820  
 2761 ATCACTCTTTTAAATATGCTGTGTTTCCCATCAAACTTGTTCAGACTTATATTTCTAA 2820  
 2821 AACCTGGAGCAAAACCAATATAATACAAACAAAGTCTGCTTCTTCAAAATCTGCTGACT 2880  
 2821 AACCTGGAGCAAAACCAATATAATACAAACAAAGTCTGCTTCTTCAAAATCTGCTGACT 2880  
 2881 CAGATATCAGTGATCTTTATAGCTTTTTCAGCCAGACATATAATGGTGACATGATTA 2940  
 2881 CAGATATCAGTGATCTTTATAGCTTTTTCAGCCAGACATATAATGGTGACATGATTA 2940  
 2941 ATGACAGTGACTATGATCTGCTGCTCCCATAGTGGGCTTTTAAATGTGATCATTCAG 3000  
 2941 ATGACAGTGACTATGATCTGCTGCTCCCATAGTGGGCTTTTAAATGTGATCATTCAG 3000  
 3001 AAAAGGACTATGTTTTCAGCTGCTTTTCAACGACTACTAGTGTATCTTTTACCTATAT 3060  
 3001 AAAAGGACTATGTTTTCAGCTGCTTTTCAACGACTACTAGTGTATCTTTTACCTATAT 3060  
 3061 TAGTGAATATCATTTAGTAACCTACTATCTTTTATCATTTTAAATGTGACTGAAACCATCCAG 3120  
 3061 TAGTGAATATCATTTAGTAACCTACTATCTTTTATCATTTTAAATGTGACTGAAACCATCCAG 3120

3121 TCTGAGTACCCCATCTTTTCAAGAAATTACTGATATAGTTTTTAAATTCAGCTGTATT 3180  
 3121 TCTGAGTACCCCATCTTTTCAAGAAATTACTGATATAGTTTTTAAATTCAGCTGTATT 3180  
 3181 TTCAAGCAGCTTTGCTTGGAAATCAATGTTACTGCAATGCCACCTTACCTTCCATGGAAA 3240  
 3181 TTCAAGCAGCTTTGCTTGGAAATCAATGTTACTGCAATGCCACCTTACCTTCCATGGAAA 3240  
 3241 ATGACAGAAATCATAAAGATCAAGCTTATATCAACTTAAACTTTTCAAGTCTTTTGGCCAT 3300  
 3241 ATGACAGAAATCATAAAGATCAAGCTTATATCAACTTAAACTTTTCAAGTCTTTTGGCCAT 3300  
 3301 CTGCATATTTGGAATTGGCAAGCTGTTGTGATATCCCTTATTTTATCATTTCTTATTT 3360  
 3301 CTGCATATTTGGAATTGGCAAGCTGTTGTGATATCCCTTATTTTATCATTTCTTATTT 3360  
 3361 TGATGCTAGGAAGCTTTATGGCAATTTTCAATATGATATATTTTATATCTGTAAGTTCC 3420  
 3361 TGATGCTAGGAAGCTTTATGGCAATTTTCAATATGATATATTTTATATCTGTAAGTTCC 3420  
 3421 TTGCTGTGTTTTTGGCTTATTTGCTTATGTTTCCATCAGTTTATTTCTGTTCACTTATATG 3480  
 3421 TTGCTGTGTTTTTGGCTTATTTGCTTATGTTTCCATCAGTTTATTTCTGTTCACTTATATG 3480  
 3481 CTTCTTCCATCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCATTTATCTATCTG 3540  
 3481 CTTCTTCCATCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCATTTATCTATCTG 3540  
 3541 TGGCAGCGTTGGCTTGTATTTGCAATCTACTGAAATTAATCTTTTATGGGATACACAAATG 3600  
 3541 TGGCAGCGTTGGCTTGTATTTGCAATCTACTGAAATTAATCTTTTATGGGATACACAAATG 3600  
 3601 CAACATTTCTTATGCTTTTGTATCTATCTATCTATCTATCTATCTATCTATCTATCTG 3660  
 3601 CAACATTTCTTATGCTTTTGTATCTATCTATCTATCTATCTATCTATCTATCTATCTG 3660  
 3661 TGATTTCTTTCATAAAGATTTCTTTGGAAGATTTGACGAAATTAATGTTGACACCTTAAATC 3720  
 3661 TGATTTCTTTCATAAAGATTTCTTTGGAAGATTTGACGAAATTAATGTTGACACCTTAAATC 3720  
 3721 CATGGATPAGCTTTCAGTACTGTTATATGCTTATCTTACCTGACGTGTGTACTGTGGATTT 3780  
 3721 CATGGATPAGCTTTCAGTACTGTTATATGCTTATCTTACCTGACGTGTGTACTGTGGATTT 3780  
 3781 TCCCTTCTTCAATATCTATGAGAAATATGAGGACAGATCAATTAAGAAAGATCCCTTTT 3840  
 3781 TCCCTTCTTCAATATCTATGAGAAATATGAGGACAGATCAATTAAGAAAGATCCCTTTT 3840  
 3841 TCAGAAACCTTTTCAACGAAAGTCTTAAATATAGAAAGCTTCCAGAACCCAGACAAATGAGG 3900  
 3841 TCAGAAACCTTTTCAACGAAAGTCTTAAATATAGAAAGCTTCCAGAACCCAGACAAATGAGG 3900  
 3901 ATGAGATGAAGATGTCAGAGCTGAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 3960  
 3901 ATGAGATGAAGATGTCAGAGCTGAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 3960  
 3961 GTTGTGAGGAGAAACCATCTTATGCTAGCAATTTGTCATATAAGAAATATGATGACAAAGA 4020  
 3961 GTTGTGAGGAGAAACCATCTTATGCTAGCAATTTGTCATATAAGAAATATGATGACAAAGA 4020  
 4021 AAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCACTTAAATATCTCTTCTTCTG 4080  
 4021 AAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCACTTAAATATCTCTTCTTCTG 4080  
 4081 TGAAAAAGGAGAGATCTTAGGACTTATTTGGTCCAAATGCTGCTGGCAAAAGCAAAATTA 4140  
 4081 TGAAAAAGGAGAGATCTTAGGACTTATTTGGTCCAAATGCTGCTGGCAAAAGCAAAATTA 4140  
 4141 TTAATATTTCTGTTGTTGATTTGAAACCACTTCAAGCCAGGTATTTTATGAGATTTAT 4200  
 4141 TTAATATTTCTGTTGTTGATTTGAAACCACTTCAAGCCAGGTATTTTATGAGATTTAT 4200

Qy	4201	CTTCAGAGACAAGTGGAAGATGATGATTCACCTGAAGTGTATGGGTATCTGTCCTCAGATAA	4260
Db	4201	CTTCAGAGACAAGTGGAAGATGATGATTCACCTGAAGTGTATGGGTATCTGTCCTCAGATAA	4260
Qy	4261	ACCTTTGTGGCCAGATACCTACATTTGCAGAAACATTTTGAATTTATGGAGCTGTCAAAG	4320
Db	4261	ACCTTTGTGGCCAGATACCTACATTTGCAGAAACATTTTGAATTTATGGAGCTGTCAAAG	4320
Qy	4321	GAATGAGTCAAGTGAATGAATGAAGAAGTCAATAAGTCGAATAAACACATGCACTGATTTAA	4380
Db	4321	GAATGAGTCAAGTGAATGAATGAAGAAGTCAATAAGTCGAATAAACACATGCACTGATTTAA	4380
Qy	4381	AAGAACATCTTCAGAACACTGTAAAGAACTACCTCGCAGAACTCAAACGAAAGTTGTGTT	4440
Db	4381	AAGAACATCTTCAGAACACTGTAAAGAACTACCTCGCAGAACTCAAACGAAAGTTGTGTT	4440
Qy	4441	TTGCTCTAAGTATGCTAGGGAATCCTCAGATTAATCTTGTAGATGAACCACTACAGGTA	4500
Db	4441	TTGCTCTAAGTATGCTAGGGAATCCTCAGATTAATCTTGTAGATGAACCACTACAGGTA	4500
Qy	4501	TGGATCCCAAGCCAAACAGACACATGTGGCGAGCAATTCGAACTGCAATTTAAACACGAA	4560
Db	4501	TGGATCCCAAGCCAAACAGACACATGTGGCGAGCAATTCGAACTGCAATTTAAACACGAA	4560
Qy	4561	AGCGGGCTGTATTTCTGACCACCTCACTATATGGAGGAGGACAGAGCTGTCTGTGATCGAG	4620
Db	4561	AGCGGGCTGTATTTCTGACCACCTCACTATATGGAGGAGGACAGAGCTGTCTGTGATCGAG	4620
Qy	4621	TAGCTATCATGGTGTCTGGGCAGTTAAGATGTATCGGAAACAGTACAACATCTAAAGAGTA	4680
Db	4621	TAGCTATCATGGTGTCTGGGCAGTTAAGATGTATCGGAAACAGTACAACATCTAAAGAGTA	4680
Qy	4681	AAATTTGAAAGAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAAACCTAGAAG	4740
Db	4681	AAATTTGAAAGAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAAACCTAGAAG	4740
Qy	4741	TAGACCGCTTCAAAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAGGAAAGTT	4800
Db	4741	TAGACCGCTTCAAAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAGGAAAGTT	4800
Qy	4801	TTTCTTCTATTTTGGCTTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCAAACTCT	4860
Db	4801	TTTCTTCTATTTTGGCTTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCAAACTCT	4860
Qy	4861	TTTTTAAGCTGGAAGAAAGCTAAACATGCTTTTGCCATTGAAGAAATATAGCTTTTCTCAAG	4920
Db	4861	TTTTTAAGCTGGAAGAAAGCTAAACATGCTTTTGCCATTGAAGAAATATAGCTTTTCTCAAG	4920
Qy	4921	CAACATTTGGAACAGGTTTTGTAGAACTCATAAGAACACAGAGGAGGAGATATAGTT	4980
Db	4921	CAACATTTGGAACAGGTTTTGTAGAACTCATAAGAACACAGAGGAGGAGATATAGTT	4980
Qy	4981	GTGGAACTTTAAACAGCACACTTTTGGTGGGAAACAAACACAGAAAGATAGATAGTATTTT	5040
Db	4981	GTGGAACTTTTAAACAGCACACTTTTGGTGGGAAACAAACACAGAAAGATAGATAGTATTTT	5040
Qy	5041	GAATTTGATTTGTTGGGTCTGCTTAATCTGGGACTTCTTTCTTTTTCACATTAATTTTAACTT	5100
Db	5041	GAATTTGATTTGTTGGGTCTGCTTAATCTGGGACTTCTTTCTTTTTCACATTAATTTTAACTT	5100
Qy	5101	TGGTTTTAAAGATTTTTTATTGGAAATGTTAACTGCGAGAACCAAGAACCGCACTTCGAAATTT	5160
Db	5101	TGGTTTTAAAGATTTTTTATTGGAAATGTTAACTGCGAGAACCAAGAACCGCACTTCGAAATTT	5160
Qy	5161	TTCTAAGCTCCTTAATTGAAATGCTGTGGTTGTGTGTTTTGCTTTTCTTTTAAATPAAACG	5220
Db	5161	TTCTAAGCTCCTTAATTGAAATGCTGTGGTTGTGTGTTTTGCTTTTCTTTTAAATPAAACG	5220
Qy	5221	TATGTAATAATTAAGTGAAGCTGCATGTTGTATTTGAAGTATATTGAACATATATAGTTTGT	5280
Db	5221	TATGTAATAATTAAGTGAAGCTGCATGTTGTATTTGAAGTATATTGAACATATATAGTTTGT	5280
Qy	5281	ATGTCATCTTTTTTCCACCATTCAGAAAACAGTGCCTTCTGAAATTTGTGATTTTAAAGGAATTGT	5340

Db	5281	ATGTCATCTTTTTCACCAATTCAGAAACAGTGCCTTCTGAATTTGTGATTTTAAAGGAATTTGT	5340
Qy	5341	AATAGAAATAGTTTTATTTTAAAGTTATCTTTAAAGTTTATGCCATCTCTTTAAATAAGTAC	5400
Db	5341	AATAGAAATAGTTTTATTTTAAAGTTATCTTTAAAGTTTATGCCATCTCTTTAAATAAGTAC	5400
Qy	5401	GTAATGTTCCAACTTAATAAATAAAAACTAATTCATACTAATGATGATGAGAAAGATACATAA	5460
Db	5401	GTAATGTTCCAACTTAATAAATAAAAACTAATTCATACTAATGATGATGAGAAAGATACATAA	5460
Qy	5461	AGCAATGTGAAAGTT 5475	
Db	5461	AGCAATGTGAAAGTT 5475	
RESULT 2			
LOCUS	AX473847	6525 bp	DNA linear PAT 09-AUG-2002
DEFINITION	Sequence 1 from Patent WO0246458.		
ACCESSION	AX473847		
VERSION	AX473847.1	GI:22208006	
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Denefle, P., Rosier-Montus, M.F., Prades, C., Arnould-Reguigne, I., Duverger, N., Allikmets, R. and Dean, M.		
TITLE	Nucleic acids of the human abca5, abca6, abca9, and abca10 genes, vectors containing such nucleic acids and uses thereof		
JOURNAL	Patent: WO 0246458-A 1 13-JUN-2002; Aventis Pharma S.A. (FR) ; The Secretary, Department of Health and Human Services (US)		
FEATURES	Location/Qualifiers		
source	1..6525		
ORIGIN	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
Query March	94.6%	Score 5181;	DB 6; Length 6525;
Best Local Similarity	99.9%	Pred. No. 0;	
Matches 5331;	Conservative 0;	Mismatches 3;	Indels 0; Gaps 0;
Qy	97	AGGTTTATTCAGAAAAACATGCTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCA	156
Db	994	AGGTTTATTCAGAAAAACATGCTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCA	1053
Qy	157	GACACCTTCTGAGAAATCTTAAATTAATGCAGAACCAAAAGAGTAGTGTTCAGG	216
Db	1054	GACACCTTCTGAGAAATCTTAAATTAATGCAGAACCAAAAGAGTAGTGTTCAGG	1113
Qy	217	AAATTCCTTTTCCACTATTTTTTTTATTTTGGTTAATATTAATAGCATGATGCATCAA	276
Db	1114	AAATTCCTTTTCCACTATTTTTTTTATTTTGGTTAATATTAATAGCATGATGCATCAA	1173
Qy	277	ATAAGAAATAGGAAGTGCCTAAATAGAACTCAATCCTATGGACAAGTTACTCTTT	336
Db	1174	ATAAGAAATAGGAAGTGCCTAAATAGAACTCAATCCTATGGACAAGTTACTCTTT	1233
Qy	337	CTAATCTAATCTTGGATATACCTCCAGTGACTAATATTACAAGCAGCATCATGCAGAAAG	396
Db	1234	CTAATCTAATCTTGGATATACCTCCAGTGACTAATATTACAAGCAGCATCATGCAGAAAG	1293
Qy	397	TGCTCTAGTCACTCTACCTGATGTCATTAATTAATCTGAAGAATATACAATATGAAGAAAA	456
Db	1294	TGCTCTAGTCACTCTACCTGATGTCATTAATTAATCTGAAGAATATACAATATGAAGAAAA	1353
Qy	457	TGTTAAATCCAGTCTCTTAAGCCGAGCACTTTGTAGTGTGGTTTTTCAAAGACTCCA	516
Db	1354	TGTTAAATCCAGTCTCTTAAGCCGAGCACTTTGTAGTGTGGTTTTTCAAAGACTCCA	1413

QY 517 TGTCTATGAACCTTCGTTTTTTTCTGATATGATTCAGTATCTTCTATTTATATGATTT 576  
 Db 1414 TGTCTATGAACCTTCGTTTTTTTCTGATATGATTCAGTATCTTCTATTTATATGATTT 1473  
 QY 577 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTGCTCAGGTTTCACAG 636  
 Db 1474 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTGCTCAGGTTTCACAG 1533  
 QY 637 TTTTACAGCATCCATAGATGCTGCCATTATACAGTTTGAAGACCAATGTTTCTCTTTGGA 696  
 Db 1534 TTTTACAGCATCCATAGATGCTGCCATTATACAGTTTGAAGACCAATGTTTCTCTTTGGA 1593  
 QY 697 AGGAGCTGGAGTCAACTAAAGCTGTTATTTAGGAGAACTGCTGTTGTGAAATAGATA 756  
 Db 1594 AGGAGCTGGAGTCAACTAAAGCTGTTATTTAGGAGAACTGCTGTTGTGAAATAGATA 1653  
 QY 757 CCTTTCCCGAGGAGTAAATTTATATACCTAGTTATAGCAATTTTCACTTTTGGTACT 816  
 Db 1654 CCTTTCCCGAGGAGTAAATTTATATACCTAGTTATAGCAATTTTCACTTTTGGTACT 1713  
 QY 817 TTTTGGCAATTCATATCGTAGCAGAAAGAAAAATAAAAGAAATTTTAAAGATAA 876  
 Db 1714 TTTTGGCAATTCATATCGTAGCAGAAAGAAAAATAAAAGAAATTTTAAAGATAA 1773  
 QY 877 TGGGACTTCATGATACCTGCTTTTGGCTTTCTCGGTTCTTCTATATACAGTTTAAATTT 936  
 Db 1774 TGGGACTTCATGATACCTGCTTTTGGCTTTCTCGGTTCTTCTATATACAGTTTAAATTT 1833  
 QY 937 TTCTATATGCTCCTCTTATATGCGAGTCATGCGACAGCTTCTTGTATTTTCCCTCAAGTA 996  
 Db 1834 TTCTATATGCTCCTCTTATATGCGAGTCATGCGACAGCTTCTTGTATTTTCCCTCAAGTA 1893  
 QY 997 GCAGCATTTGATATTTCTGCTTTTTCCTTTATGATATATCATCTGTTATTTTGGCTT 1056  
 Db 1894 GCAGCATTTGATATTTCTGCTTTTTCCTTTATGATATATCATCTGTTATTTTGGCTT 1953  
 QY 1057 TAATGCTCACACCTCTTTTAAATAAATCAAAACATGTGGGAATAGTTGAAATTTTCTTGA 1116  
 Db 1954 TAATGCTCACACCTCTTTTAAATAAATCAAAACATGTGGGAATAGTTGAAATTTTCTTGA 2013  
 QY 1117 CTGTGGCTTTTGGATTTATATGCGCTTATGATATATCCCTATAGAAAGTTTTCCTCAATCGT 1176  
 Db 2014 CTGTGGCTTTTGGATTTATATGCGCTTATGATATATCCCTATAGAAAGTTTTCCTCAATCGT 2073  
 QY 1177 TAGTGTGCTTTTTCAGTCTTTTCTGCTACTGCTACTTTTGTGATTTGATTTGCACAGTCA 1236  
 Db 2074 TAGTGTGCTTTTTCAGTCTTTTCTGCTACTGCTACTTTTGTGATTTGATTTGCACAGTCA 2133  
 QY 1237 TGCATTTTGAAGATTTTAAATGAAGTGTCTTCAATTTTCAAAATTTGACTGCGAGGCCATATC 1296  
 Db 2134 TGCATTTTGAAGATTTTAAATGAAGTGTCTTCAATTTTCAAAATTTGACTGCGAGGCCATATC 2193  
 QY 1297 CTCTAAATTTATACAAATTTATGCTTTCACCTTATAGTATATTTCTATGCTCTTGGCTG 1356  
 Db 2194 CTCTAAATTTATACAAATTTATGCTTTCACCTTATAGTATATTTCTATGCTCTTGGCTG 2253  
 QY 1357 TCTATCTTGTATCAAGTCAATCCAGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTC 1416  
 Db 2254 TCTATCTTGTATCAAGTCAATCCAGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTC 2313  
 QY 1417 TGAAGCCTTCATATTTGTTCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCCAATG 1476  
 Db 2314 TGAAGCCTTCATATTTGTTCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCCAATG 2373  
 QY 1477 TTAATGGAATATTTAGTTTATGTTAGTGAATTTATGAGCAGTTTCTTCAAGATTTGTAGGAA 1536  
 Db 2374 TTAATGGAATATTTAGTTTATGTTAGTGAATTTATGAGCAGTTTCTTCAAGATTTGTAGGAA 2433  
 QY 1537 AAGAGCCATTAAGAAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAATAATG 1596  
 Db 2434 AAGAGCCATTAAGAAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAATAATG 2493

QY 1597 AGGCTTTGAGAAATTTGTCTATTTGACATATATGAGGTCAGATTACTGCTTTACTTGGCC 1656  
 Db 2494 AGGCTTTGAGAAATTTGTCTATTTGACATATATGAGGTCAGATTACTGCTTTACTTGGCC 2553  
 QY 1657 ACAGTGGAAACAGAAAGAGTACATTTGATGAATATTTCTTTTGTGACTCTGCCACCTCTG 1716  
 Db 2554 ACAGTGGAAACAGAAAGAGTACATTTGATGAATATTTCTTTTGTGACTCTGCCACCTCTG 2613  
 QY 1717 ATGGGTTTGCATCTATATATATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 1776  
 Db 2614 ATGGGTTTGCATCTATATATATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 2673  
 QY 1777 GAAAAATGATTTGGCATTTTGTCCACAGTTAGATATACATTTGATTTTGTGACAGTAGAAG 1836  
 Db 2674 GAAAAATGATTTGGCATTTTGTCCACAGTTAGATATACATTTGATTTTGTGACAGTAGAAG 2733  
 QY 1837 AAAAAATTTATCAATTTTGGCTTCAATCAAAAGGATACACAGCCAAACAATAATACAAAGAG 1896  
 Db 2734 AAAAAATTTATCAATTTTGGCTTCAATCAAAAGGATACACAGCCAAACAATAATACAAAGAG 2793  
 QY 1897 TGCAGAGGTTTATCTAGATTTTAGACATGACAGATATCAAAAGATACCAAGCTTAAAAAT 1956  
 Db 2794 TGCAGAGGTTTATCTAGATTTTAGACATGACAGATATCAAAAGATACCAAGCTTAAAAAT 2853  
 QY 1957 TAAGTGGTGGTCAAAAAAGAGAGCTGTCATTTAGGAATTTGCTGTTCTTGGGAACCCAAAGA 2016  
 Db 2854 TAAGTGGTGGTCAAAAAAGAGAGCTGTCATTTAGGAATTTGCTGTTCTTGGGAACCCAAAGA 2913  
 QY 2017 TACTGCTGTAGTAGAACCACACAGCTGGAAATGGACCCCTGTTCTCGACATATTTGTATGA 2076  
 Db 2914 TACTGCTGTAGTAGAACCACACAGCTGGAAATGGACCCCTGTTCTCGACATATTTGTATGA 2973  
 QY 2077 ATCTTTTAAAAATACAGAAAGCCATCGGTCACAGCTGTCAGTACTCATTTTCTGAGATG 2136  
 Db 2974 ATCTTTTAAAAATACAGAAAGCCATCGGTCACAGCTGTCAGTACTCATTTTCTGAGATG 3033  
 QY 2137 AAGCTGCACATTTCTGCAGATAGAAAGCTGTGATATCAACAGGAATGCTGAAATGTTTG 2196  
 Db 3034 AAGCTGCACATTTCTGCAGATAGAAAGCTGTGATATCAACAGGAATGCTGAAATGTTTG 3093  
 QY 2197 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
 Db 3094 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153  
 QY 2257 ACAATATTTGTCACACAGAAATCTCTTCTTCTACTGTTTAAACAAACATATACCTGGAGCTA 2316  
 Db 3154 ACAATATTTGTCACACAGAAATCTCTTCTTCTACTGTTTAAACAAACATATACCTGGAGCTA 3213  
 QY 2317 CTTTATTTACACAGAAATGACCAACAACTTGTGTATAGCTTGGCTTTCAAGGACATGGACA 2376  
 Db 3214 CTTTATTTACACAGAAATGACCAACAACTTGTGTATAGCTTGGCTTTCAAGGACATGGACA 3273  
 QY 2377 AATTTTTCAGTTTGTCTGCTTCTGCTTAGCAGTCAATTTCAAAATTTGGGTGTCATTTCTTATG 2436  
 Db 3274 AATTTTTCAGTTTGTCTGCTTCTGCTTAGCAGTCAATTTCAAAATTTGGGTGTCATTTCTTATG 3333  
 QY 2437 GTGTTTCCATGACGCTTTTGGAAAGCGTATTTTAAAGCTAGAAAGTTGAGCAGAAAATG 2496  
 Db 3334 GTGTTTCCATGACGCTTTTGGAAAGCGTATTTTAAAGCTAGAAAGTTGAGCAGAAAATG 3393  
 QY 2497 ACCAGCAGATTTATAGTGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAT 2556  
 Db 3394 ACCAGCAGATTTATAGTGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAT 3453  
 QY 2557 CTTTTGATGAAATGGAACACAGAGCTTACTTATTTCTTCTGAAAACCAAGGCTTCTTAGTGA 2616  
 Db 3454 CTTTTGATGAAATGGAACACAGAGCTTACTTATTTCTTCTGAAAACCAAGGCTTCTTAGTGA 3513  
 QY 2617 GCACCATGAGCCTTTGGAAACAAACAGATGTATACATAGCAAGTTTCAATTTCTTACCT 2676  
 Db 3514 GCACCATGAGCCTTTGGAAACAAACAGATGTATACATAGCAAGTTTCAATTTCTTACCT 3573  
 QY 2677 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTGCTTTTAAATTTTTCACAG 2736







QY 577 CAAGAGCTGGCTGTTCCAAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTCACAG 636  
Db 1474 CAAGAGCTGGCTGTTCCAAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTCACAG 1533  
QY 637 TTTTCAAGCATCCATAGATGCTGCCATTATACAGTTGAAGACCAATGTTTCTCTTTGGA 696  
Db 1534 TTTTCAAGCATCCATAGATGCTGCCATTATACAGTTGAAGACCAATGTTTCTCTTTGGA 1593  
QY 697 AGGAGCTGGAGTCAACTAAAGCTGTTATATGGGAGAACTGCTGTTGTAGAAATAGATA 756  
Db 1594 AGGAGCTGGAGTCAACTAAAGCTGTTATATGGGAGAACTGCTGTTGTAGAAATAGATA 1653  
QY 757 CTTTTCCTCCGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTCACTTTTCGATACT 816  
Db 1654 CTTTTCCTCCGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTCACTTTTCGATACT 1713  
QY 817 TTTTGGCAATTCATATCTAGCAGAGAAAGAAAGAAATTAAGAAATTTTAAAGATAA 876  
Db 1714 TTTTGGCAATTCATATCTAGCAGAGAAAGAAAGAAATTAAGAAATTTTAAAGATAA 1773  
QY 877 TGGGACTTCATGATCTGCTTTTGGCTTTCTGCTGTTCTTCTATATACAAAGTTTAAATTT 936  
Db 1774 TGGGACTTCATGATCTGCTTTTGGCTTTCTGCTGTTCTTCTATATACAAAGTTTAAATTT 1833  
QY 937 TTCTTATGTCCTTCTTATGCGAGTCATTTGGACAGCTTCTTTGTATTTCTCTCAAAAGTA 996  
Db 1834 TTCTTATGTCCTTCTTATGCGAGTCATTTGGACAGCTTCTTTGTATTTCTCTCAAAAGTA 1893  
QY 997 GCAGCATTTGATATTTCTGCTTTTCTTTTCTTTATGGAATTCATCTCTATTTTTCGTT 1056  
Db 1894 GCAGCATTTGATATTTCTGCTTTTCTTTTCTTTATGGAATTCATCTCTATTTTTCGTT 1953  
QY 1057 TAAATGCTGACACCTCTTTTAAATAAATCAAAACATGTCGGAATAGTTGAAATTTTGTGA 1116  
Db 1954 TAAATGCTGACACCTCTTTTAAATAAATCAAAACATGTCGGAATAGTTGAAATTTTGTGA 2013  
QY 1117 CTGTGGCTTTTGGATTTATGTCCTTCTGCTACTGTACTTTTGTGATTTGGAATTTGCACAGTCA 1236  
Db 2014 CTGTGGCTTTTGGATTTATGTCCTTCTGCTACTGTACTTTTGTGATTTGGAATTTGCACAGTCA 2133  
QY 1237 TGCATTTAGAGATTTTAAAGAGTGTCTCATTTTCAATTTGACTGCGAGGCCATATC 1296  
Db 2134 TGCATTTAGAGATTTTAAAGAGTGTCTCATTTTCAATTTGACTGCGAGGCCATATC 2193  
QY 1297 CTCTAATTTACAAATTTATCATGCTCACACTTAATAGTATATTTCTATGCTCCTCTTGGCTG 1356  
Db 2194 CTCTAATTTACAAATTTATCATGCTCACACTTAATAGTATATTTCTATGCTCCTCTTGGCTG 2253  
QY 1357 TCTATCTTGATCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 1416  
Db 2254 TCTATCTTGATCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 2313  
QY 1417 TGAAGCTTTCATATTTGGTCAAAAGAGCAAAAGAAATTTATGAGAGTTATCAGAGGGCAATG 1476  
Db 2314 TGAAGCTTTCATATTTGGTCAAAAGAGTAAGAAATTTATGAGAGTTATCAGAGGGCAATG 2373  
QY 1477 TTAATGGAATATAGTTTATGTAATTTATGAGCCAGTTTCTTTCAGAAATTTGTAGGAA 1536  
Db 2374 TTAATGGAATATAGTTTATGTAATTTATGAGCCAGTTTCTTTCAGAAATTTGTAGGAA 2433  
QY 1537 AAGAAGCCATAAGAAATAGTGTATTTCAAGACATACAGAAAGAGGGTCAAAATGTGG 1596  
Db 2434 AAGAAGCCATAAGAAATAGTGTATTTCAAGACATACAGAAAGAGGGTCAAAATGTGG 2493  
QY 1597 AGGCTTTGAAATTTGTCATTTGACATATATGAGGGTCAGATTTCTGCCCTTACTTGGCC 1656  
Db 2494 AGGCTTTGAAATTTGTCATTTGACATATATGAGGGTCAGATTTCTGCCCTTACTTGGCC 2553  
QY 1657 ACAGTGAACAGAAAGAGTACATTTGATGAATATTTCTTTTGGAGTCTCTGCCACCTTCTG 1716

Db 2554 ACAGTGAACAGAAAGAGTACATTTGATGAATATTTCTTTGTGAGACTCTGCCACCTTCTG 2613  
QY 1717 ATGGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAA 1776  
Db 2614 ATGGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAA 2673  
QY 1777 GAAAAATGATGGCATTTCTCCACAGTTAGATATACCTTTGATGTTTTTTCAGAGTAGAAG 1836  
Db 2674 GAAAAATGATGGCATTTCTCCACAGTTAGATATACCTTTGATGTTTTTTCAGAGTAGAAG 2733  
QY 1837 AAAATTTATCAATTTTGGCTTCAATCAAAAGGATACCAAGCCAAATATTAATCAAGAAAG 1896  
Db 2734 AAAATTTATCAATTTTGGCTTCAATCAAAAGGATACCAAGCCAAATATTAATCAAGAAAG 2793  
QY 1897 TGCAGAAAGGTTTTACTAGATTTTAGACATGACACTATCAAGATACCAAGCTAAAAAT 1956  
Db 2794 TGCAGAAAGGTTTTACTAGATTTTAGACATGACACTATCAAGATACCAAGCTAAAAAT 2853  
QY 1957 TAAATGCTGGTCAAAAAAGAAAGCTGTCTCAATAGGAATTTGCTGTTCTTGGGAAACCCAAAGA 2016  
Db 2854 TAAATGCTGGTCAAAAAAGAAAGCTGTCTCAATAGGAATTTGCTGTTCTTGGGAAACCCAAAGA 2913  
QY 2017 TACTGCTGATAGTGAACCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGA 2076  
Db 2914 TACTGCTGATAGTGAACCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGA 2973  
QY 2077 ATCTTTTAAAAATACAGAAAGCAATCGGGTGCACAGTGTTCAGTACTCAATTTTCATGGATG 2136  
Db 2974 ATCTTTTAAAAATACAGAAAGCAATCGGGTGCACAGTGTTCAGTACTCAATTTTCATGGATG 3033  
QY 2137 AGCTGACATTTCTTGAGATGAGAAAGCTGTGATATCAAGGAATGCTGAAATGTTTG 2196  
Db 3034 AGCTGACATTTCTTGAGATGAGAAAGCTGTGATATCAAGGAATGCTGAAATGTTTG 3093  
QY 2197 GTTCTTCAATGTTCTTCAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
Db 3094 GTTCTTCAATGTTCTTCAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153  
QY 2257 ACAATATTTGTGCGACAGAAATCTCTTTTCTTCTAATGTTTAAACAAATATACCTGAGCTTA 2316  
Db 3154 ACAATATTTGTGCGACAGAAATCTCTTTTCTTCTAATGTTTAAACAAATATACCTGAGCTTA 3213  
QY 2317 CTTTATTTACAAAGAAATGACCAACAACTTGTGTATAGCTTTTGGCTTTCAAGGACATGGACA 2376  
Db 3214 CTTTATTTACAAAGAAATGACCAACAACTTGTGTATAGCTTTTGGCTTTCAAGGACATGGACA 3273  
QY 2377 AATTTTCAGGTTTGTCTTGGCTTAGACAGTCAATTTCAAAATTTGGGTGTCTTTATG 2436  
Db 3274 AATTTTCAGGTTTGTCTTGGCTTAGACAGTCAATTTCAAAATTTGGGTGTCTTTATG 3333  
QY 2437 GTGTTTCCATGACGCTTTTGGGAGAGCTATTTTAAAGCTAGAAATTTGAAGCAGAAATG 2496  
Db 3334 GTGTTTCCATGACGCTTTTGGGAGAGCTATTTTAAAGCTAGAAATTTGAAGCAGAAATG 3393  
QY 2497 ACCAAGCAGATTTATAGTGTATTTTACTCAGACGCTTGGAGGAAGAAATGGAATTTCAAAAT 2556  
Db 3394 ACCAAGCAGATTTATAGTGTATTTTACTCAGACGCTTGGAGGAAGAAATGGAATTTCAAAAT 3453  
QY 2557 CTTTTCATGAAATGGAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
Db 3454 CTTTTCATGAAATGGAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 3513  
QY 2617 GCACCATGAGCCTTTTGGAAACCAACAGATGTATACAAATAGCAAAAGTTTCAATTTTACCT 2676  
Db 3514 GCACCATGAGCCTTTTGGAAACCAACAGATGTATACAAATAGCAAAAGTTTCAATTTTACCT 3573  
QY 2677 TGAACCTGGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736  
Db 3574 TGAACCTGGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 3633  
QY 2737 TTCAGATTTTTATGTTTTTGGTTTCATCTCTTTTAAAAATGCTGTGTTCCCATCAAAAC 2796

3634	TT	CAGATTTT	TATGTTT	TGGTTC	ATC	AC	TCTCTTTT	TAA	AAATGCTGTG	GGTTCC	CCATCA	AAAC	3693																					
Qy	2797	TT	GTTC	CAGACT	TATAT	TTTCT	TA	AAACCT	GGAG	CAAA	CCAC	ATAA	TA	CA	AAAC	CAAGTC	2856																	
	3694	TT	GTTC	CAGACT	TATAT	TTTCT	TA	AAACCT	GGAG	CAAA	CCAC	ATAA	TA	CA	AAAC	CAAGTC	3753																	
	2857	TG	CTTCTT	CA	AAATCT	GT	CT	GA	CT	CA	GA	TAT	CAG	TAT	CT	TAT	TAG	CTTTT	CA	AAAGCC	2916													
Db	3754	TG	CTTCTT	CA	AAATCT	GT	CT	GA	CT	CA	GA	TAT	CAG	TAT	CT	TAT	TAG	CTTTT	CA	AAAGCC	3813													
	2917	AG	AA	CA	TA	AT	TGG	T	CG	CA	GAT	GA	TAA	T	AA	T	CA	CAG	T	GA	CT	AT	GT	AT	CG	T	GG	CT	CC	CA	TAG	TC	2976	
	3814	AG	AA	CA	TA	AT	TGG	T	CG	CA	GAT	GA	TAA	T	AA	T	CA	CAG	T	GA	CT	AT	GT	AT	CG	T	GG	CT	CC	CA	TAG	TC	3873	
Qy	2977	CG	CTTT	TA	AA	T	GT	GA	T	GC	AT	TC	GA	AA	GA	CA	TAT	GT	TTT	TG	CA	GT	TC	GT	TTT	CA	AA	CA	GA	TC	3036			
	3874	CG	CTTT	TA	AA	T	GT	GA	T	GC	AT	TC	GA	AA	GA	CA	TAT	GT	TTT	TG	CA	GT	TC	GT	TTT	CA	AA	CA	GA	TC	3933			
	3037	CT	AT	GG	TTT	AT	CT	TT	TA	CC	TAT	AT	TAG	T	GA	AT	CA	T	AT	TAG	T	GA	CT	AT	CA	T	AT	TT	AT	CA	TT	3096		
Db	3934	CT	AT	GG	TTT	AT	CT	TT	TA	CC	TAT	AT	TAG	T	GA	AT	CA	T	AT	TAG	T	GA	CT	AT	CA	T	AT	TT	AT	CA	TT	3993		
	3097	TA	AA	T	GT	GA	CT	GA	AA	CC	AT	CA	GA	T	CT	GG	AG	T	AC	CC	CA	T	CT	TT	TC	AA	GA	AA	TT	TA	CT	GA	TA	3156
	3994	TA	AA	T	GT	GA	CT	GA	AA	CC	AT	CA	GA	T	CT	GG	AG	T	AC	CC	CA	T	CT	TT	TC	AA	GA	AA	TT	TA	CT	GA	TA	4053
Qy	3157	TAG	TTTT	TA	AA	AT	T	GA	CT	GT	AT	TTT	CA	AG	CA	GT	TT	GT	CT	GG	AA	CA	T	AT	TT	T	GA	CT	TT	TA	CT	GA	TA	3216
	4054	TAG	TTTT	TA	AA	AT	T	GA	CT	GT	AT	TTT	CA	AG	CA	GT	TT	GT	CT	GG	AA	CA	T	AT	TT	T	GA	CT	TT	TA	CT	GA	TA	4113
	3217	TG	CA	CC	TT	TA	CT	TT	GC	CA	T	GA	AA	TC	GA	GA	AT	CA	TA	AG	AT	CA	AA	AG	CT	TT	TA	CT	TA	CT	CA	AC	3276	
Db	4114	TG	CA	CC	TT	TA	CT	TT	GC	CA	T	GA	AA	TC	GA	GA	AT	CA	TA	AG	AT	CA	AA	AG	CT	TT	TA	CT	TA	CT	CA	AC	4173	
	3277	TT	AA	AC	TTT	CA	GG	T	CT	TT	TG	CC	AT	CT	GC	AT	T	TG	GA	T	TG	CA	CA	AG	CT	TT	T	GA	T	AT	TC	3336		
	4174	TT	AA	AC	TTT	CA	GG	T	CT	TT	TG	CC	AT	CT	GC	AT	T	TG	GA	T	TG	CA	CA	AG	CT	TT	T	GA	T	AT	TC	4233		
Qy	3337	CC	TT	AT	TTT	TA	T	AT	CT	TT	AT	TT	TG	AT	GT	CT	AG	AA	CT	TA	T	TG	GC	AT	TT	CA	TT	TA	TT	AT	TC	3396		
	4234	CC	TT	AT	TTT	TA	T	AT	CT	TT	AT	TT	TG	AT	GT	CT	AG	AA	CT	TA	T	TG	GC	AT	TT	CA	TT	TA	TT	AT	TC	4293		
	3397	TA	AT	TTT	TA	T	AT	CT	TT	TA	AG	TT	CC	GT	CT	GT	GG	TTT																

QY	3877	TTCCAGAACACACCGACAAATGAGGATGAAGATGTCAAAAGCTGAAAGACTTAAAGG	3939
DB	4774	TTCCAGAACACACCGACAAATGAGGATGAAGATGTCAAAAGCTGAAAGACTTAAAGG	4833
QY	3937	TCAAAGAGCTGATGGGTGGCGTGTGTGAGGAGAAACCATCCATTTATGGTCAGCAATT	3996
DB	4834	TCAAAGAGCTGATGGGTGGCGTGTGTGAGGAGAAACCATCCATTTATGGTCAGCAATT	4893
QY	3997	TGCATAAAGAAATATGATGACAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGG	4056
DB	4894	TGCATAAAGAAATATGATGACAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGG	4953
QY	4057	CAACTAAATACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTTAGGACTATTTGGGTCCAA	4116
DB	4954	CAACTAAATACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTTAGGACTATTTGGGTCCAA	5013
QY	4117	ATGGTGTCTGCAAAAGCACAATTTAATAATTTCTGGTTGGTGATATGCAACCAACTTCAG	4176
DB	5014	ATGGTGTCTGCAAAAGCACAATTTAATAATTTCTGGTTGGTGATATGCAACCAACTTCAG	5073
QY	4177	GCCAGGTATTTTATAGGAGATTTATCTTTCCAGAGACAAGTGAAGATGATTTCACTGAAGT	4236
DB	5074	GCCAGGTATTTTATAGGAGATTTATCTTTCCAGAGACAAGTGAAGATGATTTCACTGAAGT	5133
QY	4237	GTATGGGTTACTGCTCCAGATAAACCCCTTTGTGGCCAGATATCATCATTCGACGGACATT	4296
DB	5134	GTATGGGTTACTGCTCCAGATAAACCCCTTTGTGGCCAGATATCATCATTCGACGGACATT	5193
QY	4297	TTGAAATTTTATGGAGCTGTCAAAAGGAATGAGTGCAGGTGACATGAAAGACTCATAAATC	4356
DB	5194	TTGAAATTTTATGGAGCTGTCAAAAGGAATGAGTGCAGGTGACATGAAAGACTCATAAATC	5253
QY	4357	GAATACACATGACACTGTATTTAAAGNACATCTTCAGAGACTGTATAAGAAACTACCTG	4416
DB	5254	GAATACACATGACACTGTATTTAAAGNACATCTTCAGAGACTGTATAAGAAACTACCTG	5313
QY	4417	CAGGAATCAAAAGCAAAAGTTGTGTTTTGCTCTAAGTATGCTTAGGGAATCTCTCAGATTACTT	4476
DB	5314	CAGGAATCAAAAGCAAAAGTTGTGTTTTGCTCTAAGTATGCTTAGGGAATCTCTCAGATTACTT	5373
QY	4477	TGCTAGATGAACCATCTACAGGTATGGTCCCAAGCCCAACAGACATGTGCGGAGCAA	4536
DB	5374	TGCTAGATGAACCATCTACAGGTATGGTCCCAAGCCCAACAGACATGTGCGGAGCAA	5433
QY	4537	TTCGAACTGCAATTTAAAAACAGAAAGCGGCTCTATTCTGACCACTCATATATGGAGG	4596
DB	5434	TTCGAACTGCAATTTAAAAACAGAAAGCGGCTCTATTCTGACCACTCATATATGGAGG	5493
QY	4597	AGGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGCAAGTTAAGATGTATCG	4656
DB	5494	AGGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGCAAGTTAAGATGTATCG	5553
QY	4657	GAAACAGTACAACATCTAAAGAGTAAATTTTGAAAAAGGCTACTCTTTTGGAAATTTAAATTGA	4716
DB	5554	GAAACAGTACAACATCTAAAGAGTAAATTTTGAAAAAGGCTACTCTTTTGGAAATTTAAATTGA	5613
QY	4717	AGGCTCGATAGAAACCTAGAGTAGACCGCTTCAGAGAGAAATTCAGATATATTTTCC	4776
DB	5614	AGGACTCGATAGAAACCTAGAGTAGACCGCTTCAGAGAGAAATTCAGATATATTTTCC	5673
QY	4777	CAAATGCAAGCCGTCAGAAAAAGTTTTTCTTCTATTTTGGCTTTATAAATTCCTTAAGGAAG	4836
DB	5674	CAAATGCAAGCCGTCAGAAAAAGTTTTTCTTCTATTTTGGCTTTATAAATTCCTTAAGGAAG	5733
QY	4837	ATGTTCTAGTCCCTTTCACAATCTTTTTTAAAGCTGGNAGAGCTTAAACATGCTTTTGGCCA	4896
DB	5734	ATGTTCTAGTCCCTTTCACAATCTTTTTTAAAGCTGGNAGAGCTTAAACATGCTTTTGGCCA	5793
QY	4897	TTGAAGAAATATAGCTTTTCTCAAGCAACATTGGAAACAGGTTTTTGTAGAACTCACTAAAG	4956
DB	5794	TTGAAGAAATATAGCTTTTCTCAAGCAACATTGGAAACAGGTTTTTGTAGAACTCACTAAAG	5853

```
QY 4957 AACAGAGGAGGAGAGATAATAGTTGTGGAACTTTAAACAGCACACTTTGGTGGAAAGCAA 5016
DB 5854 AACAGAGGAGGAGAGATAATAGTTGTGGAACTTTAAACAGCACACTTTGGTGGAAAGCAA 5913
QY 5017 CACAAGAGATAGAGTAGTATTTTGAATTTGATATTTGTTTGGTCTGGTCTTCTTCTTCTTCT 5076
DB 5914 CACAAGAGATAGAGTAGTATTTTGAATTTGATATTTGTTTGGTCTGGTCTTCTTCTTCTTCT 5973
QY 5077 TTTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGATTTTATTTGGAATGTAACCTGGA 5136
DB 5974 TTTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGATTTTATTTGGAATGTAACCTGGA 6033
QY 5137 GAACCAAGAACGCACTGAAATTTTCTAAGCTCTTTAATGAAATGCTGTGTTGTTG 5196
DB 6034 GAACCAAGAACGCACTGAAATTTTCTAAGCTCTTTAATGAAATGCTGTGTTGTTG 6093
QY 5197 TTTTGGCTTTCTTTAAATAAAGCAATGATTAATTAAGTGAAGCTGATGTTGTTGTTG 5256
DB 6094 TTTTGGCTTTCTTTAAATAAAGCAATGATTAATTAAGTGAAGCTGATGTTGTTGTTG 6153
QY 5257 AGTATATTGAACATATATAGTTTGTATGTCATCTTTTTCACCAATTCAGAAACAGTCTTCT 5316
DB 6154 AGTATATTGAACATATATAGTTTGTATGTCATCTTTTTCACCAATTCAGAAACAGTCTTCT 6213
QY 5317 GAATTTGTGATTTAAAGCAATGATTAATTAAGTGAAGCTGATGTTTAAAGTTTAAAGTT 5376
DB 6214 GAATTTGTGATTTAAAGCAATGATTAATTAAGTGAAGCTGATGTTTAAAGTTTAAAGTT 6273
QY 5377 TATGCCATCTTCTTTAAATAAAGCAATGATTAATTAAGTGAAGCTGATGTTTAAAGTTTAAAGTT 5430
DB 6274 TATGCCATCTTCTTTAAATAAAGCAATGATTAATTAAGTGAAGCTGATGTTTAAAGTTTAAAGTT 6327

RESULT 4
LOCUS AX392931
DEFINITION Sequence 33 from Patent WO0212340.
ACCESSION AX392931
VERSION AX392931.1 GI:19700978
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yue,H., Thornton,M., Ramkumar,J., Tang,Y.T., Azinzai,Y.,
Baughn,M.R., Yang,J., Rao,M.G., Lal,P., Wallia,N.K., Gandhi,A.R.,
Hafalia,A.J., Nguyen,D.B., Patterson,C., Elliott,V.S.,
Tribouley,C.M., Lu,D.A., Xu,Y., Reddy,R., Hernandez,R.,
Borowsky,M.L., Lo,T.P., Lu,Y., Policky,J.L., Greene,B.D.,
Sanjanwala,M.S., Raumann,B.E., Burford,N., Ison,C.H., Lee,E.A.,
Ding,L., Das,D., Kallick,D.A., Khan,P.A. and Seilhamer,J.J.
JOURNAL Patent: WO 0212340-A 33 14-FEB-2002;
FEATURES
    Location/Qualifiers
        1..6369
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
            /note="Incyte ID No: 2798241CB1"

ORIGIN
Query Match 92.9%; Score 5086; DB 6; Length 6369;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5136; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 97 AGGTTATTTCAGAAACATGTCCTCACTGCAATTAGGAGGTAGGAGTTGGAGACAGACCA 156
DB 1221 AGGTTATTTCAGAAACATGTCCTCACTGCAATTAGGAGGTAGGAGTTGGAGACAGACCA 1280
QY 157 GAACACTTCTACTGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 216
DB 1281 GAACACTTCTACTGAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1340
```

```
QY 217 AAATTCCTTTTCCACTATTTTCTTTTATTTTGGTTAAATTAATTAATTAATTAATTAATTAATTA 276
DB 1341 AAATTCCTTTTCCACTATTTTCTTTTATTTTGGTTAAATTAATTAATTAATTAATTAATTAATTA 1400
QY 277 ATAGAAATATGAGAAAGTGCTTAATATAGAACTCAATCTCTATAGCAAGTTTACTCTTT 336
DB 1401 ATAGAAATATGAGAAAGTGCTTAATATAGAACTCAATCTCTATAGCAAGTTTACTCTTT 1460
QY 337 CTAATCTAATCTTCCGATATATCTCCAGTGACTAATATTAACAAGCAGCATCATGCAAGAAAG 396
DB 1461 CTAATCTAATCTTCCGATATATCTCCAGTGACTAATATTAACAAGCAGCATCATGCAAGAAAG 1520
QY 397 TGCTCTACGATCATCTAATCTGATGTCATTAATTAATGAAAGATATACAAATGAAAGAAAGAAA 456
DB 1521 TGCTCTACGATCATCTAATCTGATGTCATTAATTAATGAAAGATATACAAATGAAAGAAAGAAA 1580
QY 457 TGCTTAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGGTGTTCCTCAAGACTCCA 516
DB 1581 TGCTTAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGGTGTTCCTCAAGACTCCA 1640
QY 517 TGCTCTATGAACCTTCGTTTCTGATATGATTCAGATATCTTCTATCTTCTATATATGGAAT 576
DB 1641 TGCTCTATGAACCTTCGTTTCTGATATGATTCAGATATCTTCTATCTTCTATATATGGAAT 1700
QY 577 CAAGAGCTGGCTGTTCAAAATCATGAGGCTGCTCAGTACTGCTGCTCAGCTTTCACAG 636
DB 1701 CAAGAGCTGGCTGTTCAAAATCATGAGGCTGCTCAGTACTGCTGCTCAGCTTTCACAG 1760
QY 637 TTTTCAAGACATCCATAGATGTCGCAATTAACAGTTCGAAGCAATGTTTCTCTTTGGA 696
DB 1761 TTTTCAAGACATCCATAGATGTCGCAATTAACAGTTCGAAGCAATGTTTCTCTTTGGA 1820
QY 697 AGGAGCTGGAGTCAACTAAAGCTGTATATATGGGAGAACTGCTGTTCTAGAAATAGATA 756
DB 1821 AGGAGCTGGAGTCAACTAAAGCTGTATATATGGGAGAACTGCTGTTCTAGAAATAGATA 1880
QY 757 CTTTTCCCGGAGGAGTAAATTTAATATACCTAGTTATAGCAATTTTCACTTTTGGATACT 816
DB 1881 CTTTTCCCGGAGGAGTAAATTTAATATACCTAGTTATAGCAATTTTCACTTTTGGATACT 1940
QY 817 TTTTGGCAATTCATATCGTAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 876
DB 1941 TTTTGGCAATTCATATCGTAGCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 2000
QY 877 TGCGACTTCATGATACTGCTTTTGGCTTTCCTGGGTTCTCTCTATATATACAGTTTAAATTT 936
DB 2001 TGCGACTTCATGATACTGCTTTTGGCTTTCCTGGGTTCTCTCTATATATACAGTTTAAATTT 2060
QY 937 TTTCTATGTCCTTCTTATGGAGTCATGCGAGAGCTTCTTTGTTATTTCTCTCAAGTA 996
DB 2061 TTTCTATGTCCTTCTTATGGAGTCATGCGAGAGCTTCTTTGTTATTTCTCTCAAGTA 2120
QY 997 GCAGCATTTGTGATATTTCTGCTTTTCTTTTATGGAATATCATCTGATTTTCTTTTCTTT 1056
DB 2121 GCAGCATTTGTGATATTTCTGCTTTTCTTTTATGGAATATCATCTGATTTTCTTTTCTTT 2180
QY 1057 TAATGCTGACACTCTTTTAAAAAATCAAAACATGTCGGATAGTTGAAATTTTGTGTA 1116
DB 2181 TAATGCTGACACTCTTTTAAAAAATCAAAACATGTCGGATAGTTGAAATTTTGTGTA 2240
QY 1117 CTGTGGCTTTTGGATTTATTTGGCTTATGATTAATCTCATAGAAAGTTTTCCTCAATCGT 1176
DB 2241 CTGTGGCTTTTGGATTTATTTGGCTTATGATTAATCTCATAGAAAGTTTTCCTCAATCGT 2300
QY 1177 TAGTGTGGCTTTTCTGCTTCTCTGTCATCTGTTTGTGATTTGGATTTGCAAGGTCA 1236
DB 2301 TAGTGTGGCTTTTCTGCTTCTCTGTCATCTGTTTGTGATTTGGATTTGCAAGGTCA 2360
QY 1237 TGCATTTTGAAGATTTTAAATGAAGGTGCTTCATTTTCAAAATTTGACTGCGAGCCCATATC 1296
DB 2361 TGCATTTTGAAGATTTTAAATGAAGGTGCTTCATTTTCAAAATTTGACTGCGAGCCCATATC 2420
QY 1297 CTCTAATTATTAACAATTAATCATGCTCACACTTAATTAATTAATTAATTAATTAATTAATTAATTA 1356
```

Db 2421 CTCTAAATTATACAAATTATCATGCTCACCTTAATAGTATATCTTAAGTCTCTTGGCTG 2480  
Qy 1357 TCATCTTTGATCAAGTCAATCCAGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTC 1416  
Db 2481 TCTATCTTTGATCAAGTCAATCCAGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTC 2540  
Qy 1417 TGAAGCCTTCATATTTGCTCAAGAGCAAAAGAAATTTATGAGGAGTTATCAGAGGGCAATG 1476  
Db 2541 TGAAGCCTTCATATTTGCTCAAGAGCAAAAGAAATTTATGAGGAGTTATCAGAGGGCAATG 2600  
Qy 1477 TTAATGAAATATTTAGTTTGTAGTGAATTTATGAGCCAGTTCTTCAGAAATTTGTAGGAA 1536  
Db 2601 TTAATGAAATATTTAGTTTGTAGTGAATTTATGAGCCAGTTCTTCAGAAATTTGTAGGAA 2660  
Qy 1537 AAGAAGCCTCAAGAAATTTAGTGGTATTTCAGAGACATACAGAAAGAGGGTGAATATGTGG 1596  
Db 2661 AAGAAGCCTCAAGAAATTTAGTGGTATTTCAGAGACATACAGAAAGAGGGTGAATATGTGG 2720  
Qy 1597 AGGCTTTGAGAAATTTGTCTATTTGACATATATGAGGGTCAGATTAAGCTTTACTTTGGCC 1656  
Db 2721 AGGCTTTGAGAAATTTGTCTATTTGACATATATGAGGGTCAGATTAAGCTTTACTTTGGCC 2780  
Qy 1657 ACAGTGAACAGGAAAGAGTACATTTGATGAATTTCTTTGTGACTCTGCCACCTTTCTG 1716  
Db 2781 ACAGTGAACAGGAAAGAGTACATTTGATGAATTTCTTTGTGACTCTGCCACCTTTCTG 2840  
Qy 1717 ATGGTTTGTATCTATATATGGAACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 1776  
Db 2841 ATGGTTTGTATCTATATATGGAACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 2900  
Qy 1777 GAAAAATGATTTGATTTGCTTCCAGTGTAGATATACATTTGATGTTTTCAGAGTAGAG 1836  
Db 2901 GAAAAATGATTTGATTTGCTTCCAGTGTAGATATACATTTGATGTTTTCAGAGTAGAG 2960  
Qy 1837 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCAAATATAATATACAGAGAG 1896  
Db 2961 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCAAATATAATATACAGAGAG 3020  
Qy 1897 TGCAGAGGTTTACTAGATTTAGACATGACAGACTATCAAGATTAACCAAGCTAAAGAT 1956  
Db 3021 TGCAGAGGTTTACTAGATTTAGACATGACAGACTATCAAGATTAACCAAGCTAAAGAT 3080  
Qy 1957 TAAGTGTGTCTCAAAAAGAAAGCTGTCTATTAGGAATTTGCTTTCTGGAAACCCAAAGA 2016  
Db 3081 TAAGTGTGTCTCAAAAAGAAAGCTGTCTATTAGGAATTTGCTTTCTGGAAACCCAAAGA 3140  
Qy 2017 TACTGTCTGTAGTGAACCAAGCTGGAATGACCCCTGTTCTCGACATATTTGATGGA 2076  
Db 3141 TACTGTCTGTAGTGAACCAAGCTGGAATGACCCCTGTTCTCGACATATTTGATGGA 3200  
Qy 2077 ATCTTTTAAAAATACAGAAAAGCCAATCGGGTGACAGTGTTCAGTACTCTCATTTTCATGGATG 2136  
Db 3201 ATCTTTTAAAAATACAGAAAAGCCAATCGGGTGACAGTGTTCAGTACTCTCATTTTCATGGATG 3260  
Qy 2137 AAGCTGACATCTTTCAGATPAGAAAAGCTGTGTATATCAAGGAATGCTGAAATGTGTTG 2196  
Db 3261 AAGCTGACATCTTTCAGATPAGAAAAGCTGTGTATATCAAGGAATGCTGAAATGTGTTG 3320  
Qy 2197 GTTCTTCAATGTCCTCAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
Db 3321 GTTCTTCAATGTCCTCAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3380  
Qy 2257 ACAAAATTTGCGCCACAGAAATCTCTTTCTTCACTGGTTAAACAAACATATACCTGGAGCTA 2316  
Db 3381 ACAAAATTTGCGCCACAGAAATCTCTTTCTTCACTGGTTAAACAAACATATACCTGGAGCTA 3440  
Qy 2317 CTTTATTAACAGATGACCAACCACTGTGTATAGCTTGCTTTCAAGGACATGGACA 2376  
Db 3441 CTTTATTAACAGATGACCAACCACTGTGTATAGCTTGCTTTCAAGGACATGGACA 3500  
Qy 2377 AAATTTTCAAGGTTGTTTCTTGGCCCTAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTATG 2436

Db 3501 AAATTTTCAAGGTTGTTTCTTGGCCCTAGACAGTCAATCAAAATTTGGGTGTCTATTTCTTATG 3560  
Qy 2437 GTGTTTCCATGACGACTTTTGAAGACGTAATTTTAAAGCTAGAAGTTGAAGCAAAATG 2496  
Db 3561 GTGTTTCCATGACGACTTTTGAAGACGTAATTTTAAAGCTAGAAGTTGAAGCAAAATG 3620  
Qy 2497 ACCAAGCAGATTTATAGTGTATTTACTCAGCAGCCACTGGAGGAAAGAAATGGAATTCAAAAT 2556  
Db 3621 ACCAAGCAGATTTATAGTGTATTTACTCAGCAGCCACTGGAGGAAAGAAATGGAATTCAAAAT 3680  
Qy 2557 CTTTTGAATGAATGGAACAGAGCTTACTTATTTCTTTCTGAAAACAGGCTTCTCTAGTGA 2616  
Db 3681 CTTTTGAATGAATGGAACAGAGCTTACTTATTTCTTTCTGAAAACAGGCTTCTCTAGTGA 3740  
Qy 2617 GCACCATGAGCCTTTTGGAAAACACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTTACCT 2676  
Db 3741 GCACCATGAGCCTTTTGGAAAACACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTTACCT 3800  
Qy 2677 TGAACGTGAAAGTAATCAATCAGTGAGTCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736  
Db 3801 TGAACGTGAAAGTAATCAATCAGTGAGTCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 3860  
Qy 2737 TTCAGATTTTATGTTTGTGTTTCATCACTCTTTTAAAGATGCTGTGCTTCCCAACAAC 2796  
Db 3861 TTCAGATTTTATGTTTGTGTTTCATCACTCTTTTAAAGATGCTGTGCTTCCCAACAAC 3920  
Qy 2797 TTGTTCCAGACTTATATTTCTTAAACCTGGAGACAAACCAATATAAATAACAAAACAATC 2856  
Db 3921 TTGTTCCAGACTTATATTTCTTAAACCTGGAGACAAACCAATATAAATAACAAAACAATC 3980  
Qy 2857 TGCTTCTTCAAAATTTCTGCTGACATCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 2916  
Db 3981 TGCTTCTTCAAAATTTCTGCTGACATCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 4040  
Qy 2917 AGAACATAATGGTGACGATGAATTAATGACAGTGACTATGTATCCGTGGCTCCCATAGT 2976  
Db 4041 AGAACATAATGGTGACGATGAATTAATGACAGTGACTATGTATCCGTGGCTCCCATAGT 4100  
Qy 2977 CGGCTTTAAATGATGATTCAGAAAAGACATGTTTTTGGACGTGTTTTCAACAGTA 3036  
Db 4101 CGGCTTTAAATGATGATTCAGAAAAGACATGTTTTTGGACGTGTTTTCAACAGTA 4160  
Qy 3037 CTATGTTTATTTCTTTTACCTATATATTAGTGAATATCAATTAGTAACTACTATCTTTATCAT 3096  
Db 4161 CTATGTTTATTTCTTTTACCTATATATTAGTGAATATCAATTAGTAACTACTATCTTTATCAT 4220  
Qy 3097 TAAATGTGACTGAAACCAATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATGATA 3156  
Db 4221 TAAATGTGACTGAAACCAATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATGATA 4280  
Qy 3157 TAGTTTTTAAAAATGAGCTGTATTTTCAAGCAGCTTTTCTTGGAAATCATTTGTTACTGCAA 3216  
Db 4281 TAGTTTTTAAAAATGAGCTGTATTTTCAAGCAGCTTTTCTTGGAAATCATTTGTTACTGCAA 4340  
Qy 3217 TGCACCTTACTTTGGCAATGGAATAATGAGAAATCAATAGATCAAAAGCTTATATCTCAAC 3276  
Db 4341 TGCACCTTACTTTGGCAATGGAATAATGAGAAATCAATAGATCAAAAGCTTATATCTCAAC 4400  
Qy 3277 TTAACCTTTCAGGCTTTTGGCAATCTCATATTTGATTTGGACAGCTGTTGTTGATATCC 3336  
Db 4401 TTAACCTTTCAGGCTTTTGGCAATCTCATATTTGATTTGGACAGCTGTTGTTGATATCC 4460  
Qy 3337 CCTTATTTTATCATCTTATTTTGTATGCTAGGAAGCTTATTTGGCAATTCATTTATGAT 3396  
Db 4461 CCTTATTTTATCATCTTATTTTGTATGCTAGGAAGCTTATTTGGCAATTCATTTATGAT 4520  
Qy 3397 TATATTTTATACGTAAAGTCCCTGCTGTGTTTTTTCCTTTTATTTGTTGTTATGTTCCAT 3456  
Db 4521 TATATTTTATACGTAAAGTCCCTGCTGTGTTTTTTCCTTTTATTTGTTGTTATGTTCCAT 4580  
Qy 3457 CAGTTATTTCTGTTCACTATATTTGCTTTCTTCCACTTTTAAAGAAATTTTAAATACCAAG 3516  
Db 4581 CAGTTATTTCTGTTCACTATATTTGCTTTCTTCCACTTTTAAAGAAATTTTAAATACCAAG 4640



3517 AATTTTGGTCATTTATCTATTCTGTGGCAGGGTTGGCTTGTTGTAATCAATCACTGAATAA 3576  
Db  
4641 AATTTTGGTCATTTATCTATTCTGTGGCAGGGTTGGCTTGTTGTAATCAATCACTGAATAA 4700  
Qy  
3577 CTTTCTTTATGGGATACACAATTCGAATCTTCTTCATTATGCGCTTTTGTATCATCTTC 3636  
Db  
4701 CTTTCTTTATGGGATACACAATTCGAATCTTCTTCATTATGCGCTTTTGTATCATCTTC 4760  
Qy  
3637 CAATCTATCCACTTCTAGGTTGGCTGATTCTTCTTCAATAAGATTCTTGGAAAGATGTAC 3696  
Db  
4761 CAATCTATCCACTTCTAGGTTGGCTGATTCTTCTTCAATAAGATTCTTGGAAAGATGTAC 4820  
Qy  
3697 GAAATAATGTGGACACCTATTAATCAATCGGATAGGCTTTGATGAGCTGTTATATCGCCTT 3756  
Db  
4821 GAAATAATGTGGACACCTATTAATCAATCGGATAGGCTTTGATGAGCTGTTATATCGCCTT 4880  
Qy  
3757 ACCTGCAGTGTACTGTGGATTTCCTCTTACAACTACTATGAGAAAAAATATGGAGCA 3816  
Db  
4881 ACCTGCAGTGTACTGTGGATTTCCTCTTACAACTACTATGAGAAAAAATATGGAGCA 4940  
Qy  
3817 GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTTAAATAATAGGAAGC 3876  
Db  
4941 GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTTAAATAATAGGAAGC 5000  
Qy  
3877 TTCCAGAACCCACGACACATGAGGATGAAGATGTCACAAAGCTGAAAGACTAAAGG 3936  
Db  
5001 TTCCAGAACCCACGACACATGAGGATGAAGATGTCACAAAGCTGAAAGACTAAAGG 5060  
Qy  
3937 TCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCATTCATTATGGTCAGCAATT 3996  
Db  
5061 TCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCATTCATTATGGTCAGCAATT 5120  
Qy  
3997 TGCATAAGAAATATCATGACAGAAAGATTTCCTTCTTCAAGAAAGTAAAGAAAGTGG 4056  
Db  
5121 TGCATAAGAAATATCATGACAGAAAGATTTCCTTCTTCAAGAAAGTAAAGAAAGTGG 5180  
Qy  
4057 CAACTAAATACATCTCTTCTGTGTGTAAGAAAGAGAGATCTTAGGACTATTGGGTCCAA 4116  
Db  
5181 CAACTAAATACATCTCTTCTGTGTGTAAGAAAGAGAGATCTTAGGACTATTGGGTCCAA 5240  
Qy  
4117 ATGGTGTGGGAAAGCAATTAATATATCTGTTGGTGGTATGTAATGAACCACTTCAG 4176  
Db  
5241 ATGGTGTGGGAAAGCAATTAATATATCTGTTGGTGGTATGTAATGAACCACTTCAG 5300  
Qy  
4177 GCCAGTATTTTAGGAGATTATCTTCAGAGACAAAGTGAAGATGATGATTCAGTGAAGT 4236  
Db  
5301 GCCAGTATTTTAGGAGATTATCTTCAGAGACAAAGTGAAGATGATGATTCAGTGAAGT 5360  
Qy  
4237 GTATGGGTTACTGCTCAGATAAACCCTTTTGTGGCCAGATACATGTCAGGAACATT 4296  
Db  
5361 GTATGGGTTACTGCTCAGATAAACCCTTTTGTGGCCAGATACATGTCAGGAACATT 5420  
Qy  
4297 TTGAAATTTATGAGCTGTCAAAGGAATGAGTGAAGTGAATGAAGAACTATAAGTC 4356  
Db  
5421 TTGAAATTTATGAGCTGTCAAAGGAATGAGTGAAGTGAATGAAGAACTATAAGTC 5480  
Qy  
4357 GAATAACATGCACTTGATTGTTAAAGAAACATTTTCAGAAAGCTGTAAAGAACTACCTG 4416  
Db  
5481 GAATAACATGCACTTGATTGTTAAAGAAACATTTTCAGAAAGCTGTAAAGAACTACCTG 5540  
Qy  
4417 CAGGAATCAAGAAAGTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTT 4476  
Db  
5541 CAGGAATCAAGAAAGTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTT 5600  
Qy  
4477 TGCTAGATGAACCATCTACAGTATGATGCCAAAGCCAAACAGCACATGTCGGAGCAA 4536  
Db  
5601 TGCTAGATGAACCATCTACAGTATGATGCCAAAGCCAAACAGCACATGTCGGAGCAA 5660  
Qy  
4537 TTGCAACTCATTTAAACAGAAAGCGGCTGCTTATCTGACCACTCACTATATGAGG 4596  
Db  
5661 TTGCAACTCATTTAAACAGAAAGCGGCTGCTTATCTGACCACTCACTATATGAGG 5720

4597 AGGCAGAGCGTCTGTGATCGAGTAGCTATCATCGTGTCTGGGCGAGTTAAGATGATCG 4656  
Db  
5721 AGGCAGAGCGTCTGTGATCGAGTAGCTATCATCGTGTCTGGGCGAGTTAAGATGATCG 5780  
Qy  
4657 GAACAGTCAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTTGA 4716  
Db  
5781 GAACAGTCAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTTGA 5840  
Qy  
4717 AGGACTGGATAGAAACCTTAGAAGTAGACCGCTTTCAAAGAGAAATTCAGTATATTTCC 4776  
Db  
5841 AGGACTGGATAGAAACCTTAGAAGTAGACCGCTTTCAAAGAGAAATTCAGTATATTTCC 5900  
Qy  
4777 CAAATGCAAGCGCTCAGAAAGTCTTCTTATTTTGGCTTATAAAATTCCTAAGGAAG 4836  
Db  
5901 CAAATGCAAGCGCTCAGAAAGTCTTCTTATTTTGGCTTATAAAATTCCTAAGGAAG 5960  
Qy  
4837 ATGTTTCAGTCCCTTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTGGCA 4896  
Db  
5961 ATGTTTCAGTCCCTTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTGGCA 6020  
Qy  
4897 TTGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTCTAGAACCTCAGTAAAG 4956  
Db  
6021 TTGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTCTAGAACCTCAGTAAAG 6080  
Qy  
4957 AACAGAGAGGAGAGATATAGTCTGGAACCTTTTAAACAGACACACTTTGGTGGGAACGAA 5016  
Db  
6081 AACAGAGAGGAGAGATATAGTCTGGAACCTTTTAAACAGACACACTTTGGTGGGAACGAA 6140  
Qy  
5017 CACAGAGATAGAGTAGTATTTTGAATTTGTTGGTCTGCTTACTGGGACTTCT 5076  
Db  
6141 CACAGAGATAGAGTAGTATTTTGAATTTGTTGGTCTGCTTACTGGGACTTCT 6200  
Qy  
5077 TTCTTTTTCACCTTAATTTTAACTTTTGGTTTAAAGAGTTTATTTGGAATGTAAGTGA 5136  
Db  
6201 TTCTTTTTCACCTTAATTTTAACTTTTGGTTTAAAGAGTTTATTTGGAATGTAAGTGA 6260  
Qy  
5137 GAACCAAGACGACCTGAAATTTTCTAAGCTCCTTAATGAAATGCTGCTGTTGTTG 5196  
Db  
6261 GAACCAAGACGACCTGAAATTTTCTAAGCTCCTTAATGAAATGCTGCTGTTGTTG 6320  
Qy  
5197 TTTTGGCTTTCTTAAATAAAACGATGATATAATAA 5233  
Db  
6321 TTTTGGCTTTCTTAAATAAAACGATGATATAATAA 6357

RESULT 5  
HSA275973

LOCUS HSA275973 5096 bp mRNA linear PRI 31-JUL-2002  
DEFINITION Homo sapiens mRNA for ATP-binding cassette protein of the (ABCA  
subfamily).

ACCESSION AJ275973  
VERSION AJ275973.3 GI:22080663  
KEYWORDS ABCA subfamily; ATP-binding cassette protein.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 Petry, F., Kotthaus, A. and Hirsch-Ernst, K.I.,  
Molecular cloning and tissue distribution of a novel ATP-binding  
cassette (ABC) transporter belonging to the subfamily ABCA  
Unpublished

## JOURNAL

REFERENCE

AUTHORS

TITLE

## JOURNAL

Submitted (02-MAR-2000) Hirsch-Ernst K.I., Department of  
Toxicology, University of Goettingen, Robert-Koch-Strasse 40,  
Lower Saxony, D-37075, GERMANY

revised by [3]

3 (bases 1 to 5096)

REFERENCE

AUTHORS

TITLE

## JOURNAL

Submitted (20-NOV-2001) Hirsch-Ernst K.I., Department of

Toxicology, University of Goettingen, Robert-Koch-Strasse 40,  
Lower Saxony, D-37075, GERMANY  
COMMENT On Aug 1, 2002 this sequence version replaced gi:17046099.  
FEATURES  
Location/Qualifiers  
1..5096  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="testis"  
66..4994  
/function="ABC transporter"  
/note="belongs to the ABCA subfamily"  
/codon\_start=1  
/product="ATP-binding cassette protein"  
/protein\_id="CAB93535.3"  
/db\_xref="gi:22080664"  
/db\_xref="GOA:Q9NY14"  
/translation="MSTAIRREVGVNRQTRLLKNLYIKRTKSSVQEIPLPFLFPLF

## CDS

WLLILSMHPNKKYEEVNIENLPMDFKTLNLIILGYTPVTNITSSIMQKVSVDHLPD  
VILTEEVNEXEMLTSSLSKPSNFVGVFKDMSYELRFPEDMIPVSIYVDSRAGCS  
KSCERAQWSSGFTVLOASIDAIIQLKTNVSLWKEHSTKAVIMGTAAVEIDTPFR  
GVLLYLVAISPGFPLAIHVAEKKIKELFKIMGLDHTAFWLSWVLLYLSLIFL  
MSLMVATASLFPFOSSSIVIFLLFPLXGLSVFFALMLTPLFKKSHVGVVEFV  
TVAFGILMILLESFPKSLVWLFSPFCHCTFVIGIAQVHLEDFNEGASFNLITAG  
PYPLIITIMLTLSIFVYLLAVLDQVIGFGLRSLVFLKPSVWSKSRNBYEL  
SEGNGNISPSEIIEPVSEFVGKEAIRISGIQKTVKKGENVPEARLNLSFDIYBQ  
ITALLGHSTGCKSLMNLICGLCPPSGFASIGHRVSEIDEMPEAKMIGICPQLDI  
HFDVLTVENLSILASIGIPANNIIQEVQVLDLDMQTIKQNAKSGSGQQRKLS  
LGIATVGPFKILLDEPTAGMDPCSRHIVNLLKRYKANRVTFVTFHMDAEDILADR  
KAVISQGLKCVGSMFLKSGYRLSMYIDRYCATESLSLVKHQI PCATILLOQD  
DQVLYLFPKDMDFSGFLSALDSHNLGVSISYGVSMITLEDVFLKLEVEAIDQD  
YSVFQOQPLEBMSKSFDEMEQSLILSETKASLVSTMSLMQOQVYIAKFHFETLK  
RESKVSRLVLLLIIFFTVOIFMFLVHSPKNAVPIKLVPDLVFLKPGDKPKHYKTS  
LLQNSRSDLSLISFTSQINWMTINDSDYVSVAHSAALNVMSHEDKYVPAVVF  
NSTWYSILFVLSNLYLHNVETIQWSTFPFQETIDIVFKLELYFQALIGI  
IVTAPPPFYAMENAEHKKAYQKLSGULPSAYWIGQAVDIPDIPFILLIMLSGL  
LAFHYGLFYFVFLAVFCILGYVPSVILFTYATSPFRKILNTKEFNSFYISVAL  
ACIALTEITFMGTATILHYAFCLIIPIYPLIAGCLISFKISWKNRKNVNTYNPW  
DRLSVATSPYQVLTFLIQVYKYGGRSIRKDPFRNLSTSKNKRKLPRPPDNE  
DEBDAERLKLGNLKMCCCEKPSIMVSNLHKEYDKKDFLSKVKVKVATKYS  
FCVKKGIELLGNAGAKSIINILVGDIEPTSGQIPLGDYSESEDDSDSLKCMGY  
CPQINLPWDDTLQHEFIEYAVKMSADMKVKSRIITHALDQHLQKLVKICPAG  
IKRKLCTFALSMLNPQITLDEPTGMDPKAKQHMWRAIRATFNKRRAAILIHTYNE  
BAEAVCRVATIMVSGQICITQVHLKSKFGYFLKIDWIEINLEVDLQREIQY  
IFPNASQESFSLAYKIPKEDVQSLSQSFPLKEEAKHAFIEEVSFSQATLEQVVF  
ELTKEBEDNSCGTLNSTLWERTQSDRVVF"

## ORIGIN

Query Match. 90.1%; Score 4935; DB 9; Length 5096;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5085; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 57 GAGCGCGTCCAGGCGTGACAGCTCTGGCGCTCGGGCCCTGAGGTTTATTTCAGAAAAACATG 116  
DB 9 GAGCGGTCAGGCGTGACAGCTCTGGCGCTCGGGCCCTGAGGTTTATTTCAGAAAAACATG 68  
QY 117 TCCACTGCAATTAGGAGGTAGGATTGGAGACAGACCAGAACACCTTCTACTGAAGAAT 176  
DB 69 TCCACTGCAATTAGGAGGTAGGATTGGAGACAGACCAGAACACCTTCTACTGAAGAAT 128  
QY 177 TACTTAATTAATGCAGAACCAAAAGAGTAGTGTCTCAGGAAATTCCTTTTCCACTATTTT 236  
DB 129 TACTTAATTAATGCAGAACCAAAAGAGTAGTGTCTCAGGAAATTCCTTTTCCACTATTTT 188  
QY 237 TTTTATTTTGGTTAATTAATTAGCATGATGATCCAAATAGAAAATATGAAGAAGTG 296  
DB 189 TTTTATTTTGGTTAATTAATTAGCATGATGATCCAAATAGAAAATATGAAGAAGTG 248  
QY 297 CCTAATATAGAACTCAATCCTATGGAAGTTTACTCTTCTTAATCTAATCTTGGATAT 356  
DB 249 CCTAATATAGAACTCAATCCTATGGAAGTTTACTCTTCTTAATCTAATCTTGGATAT 308



Db 1389 AAGAGCAAAGAAATATTAGAGAGTATTAGAGGGCAATGTTAAATGGAAATATTAGTTTT 1448  
Qy 1497 AGTGAATATTAGAGCCAGTTCTTCAGAAATTTGTAGGAAAGAGCCATAGAAATTAGT 1556  
Db 1449 AGTGAATATTAGAGCCAGTTCTTCAGAAATTTGTAGGAAAGAGCCATAGAAATTAGT 1508  
Qy 1557 GGTATTCAGAGACATACAGAAAGAGGTGAAATGTGGAGCTTTGAGAAATTTGTCA 1616  
Db 1509 GGTATTCAGAGACATACAGAAAGAGGTGAAATGTGGAGCTTTGAGAAATTTGTCA 1568  
Qy 1617 TTGACATATATGAGGGTCAGATTACTGCTTTACTTTGGCCACAGTGGAAACAGGAAAGAGT 1676  
Db 1569 TTGACATATATGAGGGTCAGATTACTGCTTTACTTTGGCCACAGTGGAAACAGGAAAGAGT 1628  
Qy 1677 ACATTGATGAATATCTTTGTGGACTCTGCCACCTCTCTGATGGGTTTGGCATATATAT 1736  
Db 1629 ACATTGATGAATATCTTTGTGGACTCTGCCACCTCTCTGATGGGTTTGGCATATATAT 1688  
Qy 1737 GGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAATGATGGCATTTGT 1796  
Db 1689 GGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAATGATGGCATTTGT 1748  
Qy 1797 CCACAGTTAGATATACATTTGATGTTTTCAGAGTAGAAGAAATTTATCAATTTGGCT 1856  
Db 1749 CCACAGTTAGATATACATTTGATGTTTTCAGAGTAGAAGAAATTTATCAATTTGGCT 1808  
Qy 1857 TCAATCAAAGGGATACAGCCCAACATATAATACAAAGATGTCAGAGGTTTACTAGAT 1916  
Db 1809 TCAATCAAAGGGATACAGCCCAACATATAATACAAAGATGTCAGAGGTTTACTAGAT 1868  
Qy 1917 TTAGACATCGACATATCAAGATAAACCAAGCTAAAGAAATTAAGTGGTGGTCAAAAAGA 1976  
Db 1869 TTAGACATCGACATATCAAGATAAACCAAGCTAAAGAAATTAAGTGGTGGTCAAAAAGA 1928  
Qy 1977 AAGCTGTCAATAGGAATGCTCTTCTGGGACCCCAAGATACCTGCTAGATGACCA 2036  
Db 1929 AAGCTGTCAATAGGAATGCTCTTCTGGGACCCCAAGATACCTGCTAGATGACCA 1988  
Qy 2037 ACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATAACAGAAA 2096  
Db 1989 ACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATAACAGAAA 2048  
Qy 2097 GCCAATCGGGTCAGAGTGTCTAGTACTCATTTTCATGGATGAGTGACATTTCTGCGAGAT 2156  
Db 2049 GCCAATCGGGTCAGAGTGTCTAGTACTCATTTTCATGGATGAGTGACATTTCTGCGAGAT 2108  
Qy 2157 AGGAAAGCTGTGATATCAAGAGGAATGCTGAAATGCTTGTGGTCTTCAATGTTCTCAAA 2216  
Db 2109 AGGAAAGCTGTGATATCAAGAGGAATGCTGAAATGCTTGTGGTCTTCAATGTTCTCAAA 2168  
Qy 2217 AGTAATGGGGATGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACAGAA 2276  
Db 2169 AGTAATGGGGATGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACAGAA 2228  
Qy 2277 TCTCTTTCTTCTACTGGTTTAAACAAATATCTCTGAGCTACTTTTATACACAGAAATGAC 2336  
Db 2229 TCTCTTTCTTCTACTGGTTTAAACAAATATCTCTGAGCTACTTTTATACACAGAAATGAC 2288  
Qy 2337 CAACAACTTGTGTATAGCTTTCAGGACATGGAACAAATTTCAAGTTTGTGTTTCT 2396  
Db 2289 CAACAACTTGTGTATAGCTTTCAGGACATGGAACAAATTTCAAGTTTGTGTTTCT 2348  
Qy 2397 GCCCTAGACAGTCATTCAAATTTGGGTGTCATTCTTATGTTGTTTCCATGACGACTTG 2456  
Db 2349 GCCCTAGACAGTCATTCAAATTTGGGTGTCATTCTTATGTTGTTTCCATGACGACTTG 2408  
Qy 2457 GAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACAGAGATTTAGTGTA 2516  
Db 2409 GAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACAGAGATTTAGTGTA 2468  
Qy 2517 TTTTACTCAGACGCACTGGAGAGAAATGGATTCAAATCTTTTGTGTAATGGACAG 2576

Db 2469 TTTTACTCAGACGCACTGGAGGAAGAAATGGATTCAAATCTTTTGTGAAATGGAAACAG 2528  
Qy 2577 AGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAGCAACCATAGGCTTTTGGAAA 2636  
Db 2529 AGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAGCAACCATAGGCTTTTGGAAA 2588  
Qy 2637 CAACAGATGTATCAATAGCAAGTTTCATTTCTTACCTTGAAGCGTGAAAGTAAATCA 2696  
Db 2589 CAACAGATGTATCAATAGCAAGTTTCATTTCTTACCTTGAAGCGTGAAAGTAAATCA 2648  
Qy 2697 GTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAGTTTCAGATTTTTTGTG 2756  
Db 2649 GTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAGTTTCAGATTTTTTGTG 2708  
Qy 2757 GTTTCATCATCTTTTAAAAATGCTGTGGTTCGCATCAAACTTGTTCACAGCTTATATTT 2816  
Db 2709 GTTTCATCATCTTTTAAAAATGCTGTGGTTCGCATCAAACTTGTTCACAGCTTATATTT 2768  
Qy 2817 CTAAAACTGGAGACAAACCAATAAATACAAAAAAGTCTGCTTCTTCAAAATTTCTGCT 2876  
Db 2769 CTAAAACTGGAGACAAACCAATAAATACAAAAAAGTCTGCTTCTTCAAAATTTCTGCT 2828  
Qy 2877 GACTCAGATATCAGTGATCTTATAGCTTTTTCAGAGCCAGAACATATGTCAGCATG 2936  
Db 2829 GACTCAGATATCAGTGATCTTATAGCTTTTTCAGAGCCAGAACATATGTCAGCATG 2888  
Qy 2937 ATTAAATGACAGTGACCTATGATCCGTGGCTCCCATAGTGGGCTTTAAATGTGATGCAT 2996  
Db 2889 ATTAAATGACAGTGACCTATGATCCGTGGCTCCCATAGTGGGCTTTAAATGTGATGCAT 2948  
Qy 2997 TCAGAAAAAGCATATGTTTTTTCAGAGCTGTTTTTTCAGAGTACTATGTTTTTCTTTACCT 3056  
Db 2949 TCAGAAAAAGCATATGTTTTTTCAGAGCTGTTTTTTCAGAGTACTATGTTTTTCTTTACCT 3008  
Qy 3057 ATATTAGTGAATATCAATAGTAACTACTATCTTTTCAAGAAATTTCTGATATAGTTTTTAAATGTGAGCTG 3116  
Db 3009 ATATTAGTGAATATCAATAGTAACTACTATCTTTTCAAGAAATTTCTGATATAGTTTTTAAATGTGAGCTG 3068  
Qy 3117 CAGATCTGAGTACCCCATCTTTTCAAGAAATTTCTGATATAGTTTTTAAATGTGAGCTG 3176  
Db 3069 CAGATCTGAGTACCCCATCTTTTCAAGAAATTTCTGATATAGTTTTTAAATGTGAGCTG 3128  
Qy 3177 TATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACTTACTTTGCGCATG 3236  
Db 3129 TATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACTTACTTTGCGCATG 3188  
Qy 3237 GAAAAATGCGAGAAATCATAGATCAAAGCTTATACCTTAACTTTAACTTTCAAGTCTTTTG 3296  
Db 3189 GAAAAATGCGAGAAATCATAGATCAAAGCTTATACCTTAACTTTAACTTTCAAGTCTTTTG 3248  
Qy 3297 CCATCTGCATATTTGGATGGAACAGCTGTTGTGATATCCCTTATTTTTTATCATTTCT 3356  
Db 3249 CCATCTGCATATTTGGATGGAACAGCTGTTGTGATATCCCTTATTTTTTATCATTTCT 3308  
Qy 3357 ATTTTGTATGCTAGAAAGCTTATTTGGCATTTCAATTTGATATATTTTTTATCTGTAAG 3416  
Db 3309 ATTTTGTATGCTAGAAAGCTTATTTGGCATTTCAATTTGATATATTTTTTATCTGTAAG 3368  
Qy 3417 TTCTTGTCTGTGGTTTTTTTGGCTTATTTGTTATGTTTCCATGAGTATTTCTGTTCACTTAT 3476  
Db 3369 TTCTTGTCTGTGGTTTTTTTGGCTTATTTGTTATGTTTCCATGAGTATTTCTGTTCACTTAT 3428  
Qy 3477 ATTTGCTTCTTTTCACTTTTAAAGAAATTTTTTAAATACCAAAGAAATTTTGGTCAATTTCTAT 3536  
Db 3429 ATTTGCTTCTTTTCACTTTTAAAGAAATTTTTTAAATACCAAAGAAATTTTGGTCAATTTCTAT 3488  
Qy 3537 TCTGTGGCAGCGTTGGCTGTATGCAATCACTGAAATTAACCTTTCTTTATGGGATACACA 3596  
Db 3489 TCTGTGGCAGCGTTGGCTGTATGCAATCACTGAAATTAACCTTTCTTTATGGGATACACA 3548  
Qy 3597 ATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTTCCCAATCTATCCACTTTCTAGCT 3656  
Db 3549 ATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTTCCCAATCTATCCACTTTCTAGCT 3608

QY 3657 TGCCTGATTTCTTTTCATAAAGATTTCTTGGAAAGATGTACGAAAAAATGTGGACACCTAT 3716  
 Db 3609 TGCCTGATTTCTTTTCATAAAGATTTCTTGGAAAGATGTACGAAAAAATGTGGACACCTAT 3668  
 QY 3717 AATCAATGGATAGGCTTTTCACTAGCTGTATATCGCCTTACCTCGAGTGTACTGTGG 3776  
 Db 3669 AATCAATGGATAGGCTTTTCACTAGCTGTATATCGCCTTACCTCGAGTGTACTGTGG 3728  
 QY 3777 ATTTTCCTCTTCAATATCATATATGAGAAAAAATATGAGGCGAGATCAATTAAGAAAAAGATCCC 3836  
 Db 3729 ATTTTCCTCTTCAATATCATATATGAGAAAAAATATGAGGCGAGATCAATTAAGAAAAAGATCCC 3788  
 QY 3837 TTTTTCAGAAACCTTTCAACGAAGTCTTAAATATGGAAGTCTTCAAGACCCAGACCAAT 3896  
 Db 3789 TTTTTCAGAAACCTTTCAACGAAGTCTTAAATATGGAAGTCTTCAAGACCCAGACCAAT 3848  
 QY 3897 GAGGATGAAGATGAAGATGTCAAAAGCTGAAAGCTTAAAGGTCAAAAGCTGATGGGTGC 3956  
 Db 3849 GAGGATGAAGATGAAGATGTCAAAAGCTGAAAGCTTAAAGGTCAAAAGCTGATGGGTGC 3908  
 QY 3957 CAGTGTGTGAGAGAAACCTTCAATATATGCTCAGCAATTTGCTAATAAGATATGATGAC 4016  
 Db 3909 CAGTGTGTGAGAGAAACCTTCAATATATGCTCAGCAATTTGCTAATAAGATATGATGAC 3968  
 QY 4017 AAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTTAAATATCATCTCTTC 4076  
 Db 3969 AAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTTAAATATCATCTCTTC 4028  
 QY 4077 TGTGTGAAAAAGAGAGATCTTAGGACTATTTGGGTCCAAATGGTGGTGGCAAAAGACACA 4136  
 Db 4029 TGTGTGAAAAAGAGAGATCTTAGGACTATTTGGGTCCAAATGGTGGTGGCAAAAGACACA 4088  
 QY 4137 ATTTATTAATTTCTGGTGTGATATGAAACCACTTCAGGCCAGGATATTTTATGAGGAT 4196  
 Db 4089 ATTTATTAATTTCTGGTGTGATATGAAACCACTTCAGGCCAGGATATTTTATGAGGAT 4148  
 QY 4197 TATTTTCAGAGCAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTACTGCTCTCAG 4256  
 Db 4149 TATTTTCAGAGTCAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTACTGCTCTCAG 4208  
 QY 4257 ATAAACCTTTTGGCCAGTACTACATTTGAGGAGCACTTTTGAATTTATGAGGCTGTC 4316  
 Db 4209 ATAAACCTTTTGGCCAGTACTACATTTGAGGAGCACTTTTGAATTTATGAGGCTGTC 4268  
 QY 4317 AAGGAATGAGTCAAGTGAATGAAGAAAGTCAATGCTGAATGAACACATGCACTCTGAT 4376  
 Db 4269 AAGGAATGAGTCAAGTGAATGAAGAAAGTCAATGCTGAATGAACACATGCACTCTGAT 4328  
 QY 4377 TTAAGAAACATCTTCAGAGACTGTAAAGAAACTACCTGAGGAACTCAACGAAAGTTG 4436  
 Db 4329 TTAAGAAACATCTTCAGAGACTGTAAAGAAACTACCTGAGGAACTCAACGAAAGTTG 4388  
 QY 4437 TGTTTTGTCTTAAGTATGCTAGGAAATCTCTCAGATTTACTTTGCTAGATGAACCATCTACA 4496  
 Db 4389 TGTTTTGTCTTAAGTATGCTAGGAAATCTCTCAGATTTACTTTGCTAGATGAACCATCTACA 4448  
 QY 4497 GGTATGATATCCAAAGCAACAGACATGTGGGAGCAATTCGAACTGCAATTTAAAAAC 4556  
 Db 4449 GGTATGATATCCAAAGCAACAGACATGTGGGAGCAATTCGAACTGCAATTTAAAAAC 4508  
 QY 4557 AGAAAGCGGGTGTCTTCTGACCACTTCACTATATGAGGAGGAGGAGGCTGTCTGTGAT 4616  
 Db 4509 AGAAAGCGGGTGTCTTCTGACCACTTCACTATATGAGGAGGAGGAGGCTGTCTGTGAT 4568  
 QY 4617 CAGTGTATCATGGTGTCTGGCAGTTAAGATGTATCGGAAACAGTACCAACTCTAAAG 4676  
 Db 4569 CAGTGTATCATGGTGTCTGGCAGTTAAGATGTATCGGAAACAGTACCAACTCTAAAG 4628  
 QY 4677 AGTAAATTTGGAAGGCTCTTTTGGAAATTAATTTGAGGAGCTGGATAGAAAACTTA 4736  
 Db 4629 AGTAAATTTGGAAGGCTCTTTTGGAAATTAATTTGAGGAGCTGGATAGAAAACTTA 4688

QY 4737 GAAGTAGACCGCCTTCAAGAGAAAAATTCAGTATATTTTCCAAATCAAGCCGTCAGGAA 4796  
 Db 4689 GAAGTAGACCGCCTTCAAGAGAAAAATTCAGTATATTTTCCAAATCAAGCCGTCAGGAA 4748  
 QY 4797 AGTTTTCCTCTTCTTATTTGGCTTATAAAATTCCTTAAGGAGATGTTTCACTCCCTTTCACAA 4856  
 Db 4749 AGTTTTCCTCTTCTTATTTGGCTTATAAAATTCCTTAAGGAGATGTTTCACTCCCTTTCACAA 4808  
 QY 4857 TCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCCATTTGAAGAATAATAGCTTTTCT 4916  
 Db 4809 TCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCCATTTGAAGAATAATAGCTTTTCT 4868  
 QY 4917 CAAGCAACATTTGAAACAGCTTTTGTAGAACTCACTAAGAAACAAAGAGGAGAGATAT 4976  
 Db 4869 CAAGCAACATTTGAAACAGCTTTTGTAGAACTCACTAAGAAACAAAGAGGAGAGATAT 4928  
 QY 4977 AGTTGTGGAATTTTAAACAGCACACTTTTGTGGGAAACGAAACACAAAGATAGAGTAGTA 5036  
 Db 4929 AGTTGTGGAATTTTAAACAGCACACTTTTGTGGGAAACGAAACACAAAGATAGAGTAGTA 4988  
 QY 5037 TTTTGAATTTGTATTTGGTGTCTGCTTACTGGGACTTCTTTTCTTTTCACTTAATTTTA 5096  
 Db 4989 TTTTGAATTTGTATTTGGTGTCTGCTTACTGGGACTTCTTTTCTTTTCACTTAATTTTA 5048  
 QY 5097 ACTTTGGTTTAAAGCTTTTATTTGGAATGGTAACTGGAGAACCAAG 5144  
 Db 5049 ACTTTGGTTTAAAGCTTTTATTTGGAATGGTAACTGGAGAACCAAG 5096

RESULT 6  
 LOCUS AX417828 5262 bp DNA linear PAT 18-JUN-2002  
 DEFINITION Sequence 5 from Patent WO0231147.  
 ACCESSION AX417828  
 VERSION AX417828.1 GI:21522944  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 Hu, Y. and Nepomnichy, B.  
 Human transporter proteins and polynucleotides encoding the same  
 Patent: WO 0231147-A 5 18-APR-2002;  
 LEXICON GENETICS INC (US)  
 FEATURES  
 Location/Qualifiers  
 1..5262  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 85.2%; Score 4664; DB 6; Length 5262;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 5014; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 97 AGGTTTATTCAGAAACATGTCCTGCAATTTAGGAGGTAGAGTTTGGACAGACCA 156  
 Db 238 AGGTTTATTCAGAAACATGTCCTGCAATTTAGGAGGTAGAGTTTGGACAGACCA 297  
 QY 157 GAACACTTCTACTGGAAGATTAATTAATTAATGCAAGAACCAAAAGAGTAGTTTCAGG 216  
 Db 298 GAACACTTCTACTGGAAGATTAATTAATTAATGCAAGAACCAAAAGAGTAGTTTCAGG 357  
 QY 217 AAATCTTTTCCACATTTTATTTTATTTGTTGTTTATTAATTAATGATGATGATCCAA 276  
 Db 358 AAATCTTTTCCACATTTTATTTTATTTGTTGTTTATTAATTAATGATGATGATCCAA 417  
 QY 277 ATAAGAAATATGAAGAGTGCCTTAATATAGAACTCAATCTATGCAAGAGTTTACTCTTT 336  
 Db 418 ATAAGAAATATGAAGAGTGCCTTAATATAGAACTCAATCTATGCAAGAGTTTACTCTTT 477  
 QY 337 CTAATCTAATCTTGTGATATCTCCAGTGTACTTAATATTAAGAGCATCATGTCAGAAAG 396

Db	478	CTAATCTAAATTCCTTGGATATATCTCCAGTGATTAATATATACAGCAGATCATATGCAGAAAG	537
Qy	397	TGCTACTGATCATCTACTCTGATGTCAATAATCTCTGAAGAATATACAAATGAAAGAAA	456
Db	538	TGTCTACTGATCATCTACTCTGATGTCAATAATCTCTGAAGAATATACAAATGAAAGAAA	597
Qy	457	TGTTAAACATCCAGTCTCTCTAAGCGGAGCACTTTGTAGTGTGGTTTCCAAAGACTCCA	516
Db	598	TGTTAAACATCCAGTCTCTCTAAGCGGAGCACTTTGTAGTGTGGTTTCCAAAGACTCCA	657
Qy	517	TGTCCTATGAACCTTCGTTTTTTCCTGATATGATCCAGTATCTCTATTTATATGGATT	576
Db	658	TGTCCTATGAACCTTCGTTTTTTCCTGATATGATCCAGTATCTCTATTTATATGGATT	717
Qy	577	CAAGAGCTGGCTGTTCAAAATCATGTAGGCTGCTCAGTACTGTGCTCTCAGGTTTCACAG	636
Db	718	CAAGAGCTGGCTGTTCAAAATCATGTAGGCTGCTCAGTACTGTGCTCTCAGGTTTCACAG	777
Qy	637	TTTTACAAGCATCCATAGATGCTGCCAATATACAGTTGAAGACCAATGTTTCTCTTTGGA	696
Db	778	TTTTACAAGCATCCATAGATGCTGCCAATATACAGTTGAAGACCAATGTTTCTCTTTGGA	837
Qy	697	AGGAGCTGGAGTCAACTAAAGCTGTATATATGGAGAAACCTGCTGTGTAGAAATAGATA	756
Db	838	AGGAGCTGGAGTCAACTAAAGCTGTATATATGGAGAAACCTGCTGTGTAGAAATAGATA	897
Qy	757	CCTTTCOCAGGAGTAAATTTAATAATACCTAGTTATAGCAATTTCCACTTTTGGATPACT	816
Db	898	CCTTTCOCAGGAGTAAATTTAATAATACCTAGTTATAGCAATTTCCACTTTTGGATPACT	957
Qy	817	TTTTGGCAATTCATATCGTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAA	876
Db	958	TTTTGGCAATTCATATCGTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAA	1017
Qy	877	TGGACATTCATGATCTGCCTTTTGGCTTTTCCCTGGGTTCTCTATATACAGTTTAATTT	936
Db	1018	TGGACATTCATGATCTGCCTTTTGGCTTTTCCCTGGGTTCTCTATATACAGTTTAATTT	1077
Qy	937	TTCTTATGTGCCCTTCTTATGSCAGTCAATGGCAGAGCTTCTTTGTTATTTCCCTCAAAGTA	996
Db	1078	TTCTTATGTGCCCTTCTTATGSCAGTCAATGGCAGAGCTTCTTTGTTATTTCCCTCAAAGTA	1137
Qy	997	GCAGATTGTGATATTTCTGCTTTTTTTCCTTTATGAAATATCACTGTATTTTTTGCTT	1056
Db	1138	GCAGATTGTGATATTTCTGCTTTTTTTCCTTTATGAAATATCACTGTATTTTTTGCTT	1197
Qy	1057	TAAATGCTGACACCTCTTTTAAAAAATCAAAACATGTGGGAATAGTTGAAATTTTTTGCTA	1116
Db	1198	TAAATGCTGACACCTCTTTTAAAAAATCAAAACATGTGGGAATAGTTGAAATTTTTTGCTA	1257
Qy	1117	CTGTGGCTTTTGGATTTATTTGGCCCTTATGATAATCTCTCATAGAAAGTTTCCCAATCGT	1176
Db	1258	CTGTGGCTTTTGGATTTATTTGGCCCTTATGATAATCTCTCATAGAAAGTTTCCCAATCGT	1317
Qy	1177	TAGTGTGGCTTTTCAGTCCCTTCTGTCACTGTACTTTTGTGATTTGATTTGCAAGTCA	1236
Db	1318	TAGTGTGGCTTTTCAGTCCCTTCTGTCACTGTACTTTTGTGATTTGATTTGCAAGTCA	1377
Qy	1237	TGCATTTAGAAGATTTTAAATGAAGGTGCTTCATTTTCAAATTTGACTGCGAGGCCATATC	1296
Db	1378	TGCATTTAGAAGATTTTAAATGAAGGTGCTTCATTTTCAAATTTGACTGCGAGGCCATATC	1437
Qy	1297	CTCTAAATTTATPACAAATATCATGCTCACATTAATAGTATATTTCTATGCTCTTGGCTG	1356
Db	1438	CTCTAAATTTATPACAAATATCATGCTCACATTAATAGTATATTTCTATGCTCTTGGCTG	1497
Qy	1357	TCTATCTTGATCAAGTCAATTTCCAGGGGAAATTTGGCTTACGGAGATCATCTTTATTTTC	1416
Db	1498	TCTATCTTGATCAAGTCAATTTCCAGGGGAAATTTGGCTTACGGAGATCATCTTTATTTTC	1557
Qy	1417	TGAAGCCTTTCATATTTGGTCAAAGAGCAAAAAGAAATATAGGAGCTATACAGAGGCCAATG	1476

QY	2557	CTTTTGATGAATGGAAACAGAGCTTACTTATTTCTTCTGAAACCAAGCCTTCTCTAGTGA	2616
Db	2698	CTTTTGATGAATGGAAACAGAGCTTACTTATTTCTTCTGAAACCAAGCCTTCTCTAGTGA	2757
QY	2617	GCACCATGAGCCTTTGAAACCAACAGATGATACAAATAGCAAGTTTCATTTCTTACCT	2676
Db	2758	GCACCATGAGCCTTTGAAACCAACAGATGATACAAATAGCAAGTTTCATTTCTTACCT	2817
QY	2677	TGAAACGTGAAAGTAATACAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG	2736
Db	2818	TGAAACGTGAAAGTAATACAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG	2877
QY	2737	TTCAAGATTTTATGTTTTTGGTTTCATCACTCTTTTAAAAATGCTGGTTCCCATCAAAAC	2796
Db	2878	TTCAAGATTTTATGTTTTTGGTTTCATCACTCTTTTAAAAATGCTGGTTCCCATCAAAAC	2937
QY	2797	TTGTTCCAGACTTATATTTCTTAAACCTGGAGACAAACACATAAATACAAACAAAGTC	2856
Db	2938	TTGTTCCAGACTTATATTTCTTAAACCTGGAGACAAACACATAAATACAAACAAAGTC	2997
QY	2857	TGCTTCTTCAAAATTCGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTTCACAAGCC	2916
Db	2998	TGCTTCTTCAAAATTCGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTTCACAAGCC	3057
QY	2917	AGAACATAAGGTGACCAATGATTAATGACAGTGACTATGATCGGTGGCTCCCATAGTG	2976
Db	3058	AGAACATAAGGTGACCAATGATTAATGACAGTGACTATGATCGGTGGCTCCCATAGTG	3117
QY	2977	CGGCTTTAAATGTGATGCATTCAGAAAAGGACTATGTTTTTGAGCTGTTTTTCAACAGTA	3036
Db	3118	CGGCTTTAAATGTGATGCATTCAGAAAAGGACTATGTTTTTGAGCTGTTTTTCAACAGTA	3177
QY	3037	CTATGGTTTATCTTTTACCTATATAGTGAATATCATTAGTAATCTTATCTTTTATCAT	3096
Db	3178	CTATGGTTTATCTTTTACCTATATAGTGAATATCATTAGTAATCTTATCTTTTATCAT	3237
QY	3097	TAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATPACTGATA	3156
Db	3238	TAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATPACTGATA	3297
QY	3157	TAGTTTTTAAATGTAGCTGATTTTCAAGACGCTTTGCTTGGAAATCATTTGTTACTGCAA	3216
Db	3298	TAGTTTTTAAATGTAGCTGATTTTCAAGACGCTTTGCTTGGAAATCATTTGTTACTGCAA	3357
QY	3217	TGCCACCTTACTTTGCCATG3AAATGACAGATCATATAGATCAAGCTTATPACTCAAC	3276
Db	3358	TGCCACCTTACTTTGCCATG3AAATGACAGATCATATAGATCAAGCTTATPACTCAAC	3417
QY	3277	TTAAACTTTTCAGGTCTTTTGCCATCTGCATATTTGGATTTGGAACGCTGTTGTTGATATCC	3336
Db	3418	TTAAACTTTTCAGGTCTTTTGCCATCTGCATATTTGGATTTGGAACGCTGTTGTTGATATCC	3477
QY	3337	CCTTATTTTTTATCATTTCTTATTTGATGCTAGGAAGCTTATTTGGCATTTTATATGGAT	3396
Db	3478	CCTTATTTTTTATCATTTCTTATTTGATGCTAGGAAGCTTATTTGGCATTTTATATGGAT	3537
QY	3397	TATATTTTTTATCTGTAAGTTCCTTCTGCTGGTTTTTTTGGCTTATTTGGTTATGTTCCAT	3456
Db	3538	TATATTTTTTATCTGTAAGTTCCTTCTGCTGGTTTTTTTGGCTTATTTGGTTATGTTCCAT	3597
QY	3457	CAGTTATCTGTTCACTTATATTTGCTTTTCCACCTTTAAGAAAAATTTTAAATACCAAG	3516
Db	3598	CAGTTATCTGTTCACTTATATTTGCTTTTCCACCTTTAAGAAAAATTTTAAATACCAAG	3657
QY	3517	AATTTTGGTCAATTTATCTATTTCTGTGGCAGCTTTGGCTTGTATTTGCAATCACTGAAATAA	3576
Db	3658	AATTTTGGTCAATTTATCTATTTCTGTGGCAGCTTTGGCTTGTATTTGCAATCACTGAAATAA	3717
QY	3577	CTTTCTTTTATGGGATACAAATGCAACTATTTCTTATATGCTTTTGTATCATCATTC	3636
Db	3718	CTTTCTTTTATGGGATACAAATGCAACTATTTCTTATATGCTTTTGTATCATCATTC	3777

QY	3637	CAATCTATCCACTTCTAGTGTGCTGATTTCTTTTATATAAGATTTCTTTGGAAAGATGTAC	3696
Db	3778	CAATCTATCCACTTCTAGTGTGCTGATTTCTTTTATATAAGATTTCTTTGGAAAGATGTAC	3837
QY	3697	GAAAAAATGTGACACCTTATTAATCCATGGGATAGGCTTTTCAGTAGCTGTATATATGCGCTT	3756
Db	3838	GAAAAAATGTGACACCTTATTAATCCATGGGATAGGCTTTTCAGTAGCTGTATATATGCGCTT	3897
QY	3757	ACCTCAGTGTGTACTGTGGGATTTTCTCTTACATACTATGAGAAAAAATATAGAGCA	3816
Db	3898	ACCTCAGTGTGTACTGTGGGATTTTCTCTTACATACTATGAGAAAAAATATAGAGCA	3957
QY	3817	GATCAATAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAGTCTAAAAATAGGAAGC	3876
Db	3958	GATCAATAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAGTCTAAAAATAGGAAGC	4017
QY	3877	TTCCAGAACCCACGACACATGAGGATGAAGATGAAGATGTCAAGCTCAAGAGCTAAAGG	3936
Db	4018	TTCCAGAACCCACGACACATGAGGATGAAGATGAAGATGTCAAGCTCAAGAGCTAAAGG	4077
QY	3937	TCAAAAGAGCTGATGGGTTGCCAGTGTGTTGAGGAGAAAACCATCCATTTATGTCAGCAATT	3996
Db	4078	TCAAAAGAGCTGATGGGTTGCCAGTGTGTTGAGGAGAAAACCATCCATTTATGTCAGCAATT	4137
QY	3997	TGCATAAAGATATGATGACAAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAACTGG	4056
Db	4138	TGCATAAAGATATGATGACAAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAACTGG	4197
QY	4057	CAACTAAAATACATCTCTTTCTGTGAAAAAAGAGAGATCTTAGGACTATTTGGGTCCAA	4116
Db	4198	CAACTAAAATACATCTCTTTCTGTGAAAAAAGAGAGATCTTAGGACTATTTGGGTCCAA	4257
QY	4117	ATGGTCTGGGAAAAGCAAAATTAATTAATTTCTGGTTGGTGATTTGAAACCACTTCAG	4176
Db	4258	ATGGTCTGGGAAAAGCAAAATTAATTAATTTCTGGTTGGTGATTTGAAACCACTTCAG	4317
QY	4177	GCAGGTATTTTATAGAGATTTATCTTCAGAGACAAGTGAAGATGATGATTTCACTGAAGT	4236
Db	4318	GCAGGTATTTTATAGAGATTTATCTTCAGAGACAAGTGAAGATGATGATTTCACTGAAGT	4377
QY	4237	GTATGGGTTACTGTCTCAGATAAACCCCTTTGTGGCCAGATATCTACATTTGCGAGAACAT	4296
Db	4378	GTATGGGTTACTGTCTCAGATAAACCCCTTTGTGGCCAGATATCTACATTTGCGAGAACAT	4437
QY	4297	TTGAAATTTATGGAGCTGTCAAGAGATGAGTGAAGTGACATGAAAGAGTCTATAAGTC	4356
Db	4438	TTGAAATTTATGGAGCTGTCAAGAGATGAGTGAAGTGACATGAAAGAGTCTATAAGTC	4497
QY	4357	GAATAACACATGCACTTGATTTTAAAGAAACATCTTCAGAGAACCTGTAAAGAACTTACCTG	4416
Db	4498	GAATAACACATGCACTTGATTTTAAAGAAACATCTTCAGAGAACCTGTAAAGAACTTACCTG	4557
QY	4417	CAGGATCAAGAGAAAGTTGTTTTGCTCTAAGTATGCTAGGGAATCCTCAGATTACTT	4476
Db	4558	CAGGATCAAGAGAAAGTTGTTTTGCTCTAAGTATGCTAGGGAATCCTCAGATTACTT	4617
QY	4477	TGCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAAACAGCATGTGGGAGCAA	4536
Db	4618	TGCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAAACAGCATGTGGGAGCAA	4677
QY	4537	TTGCAACTGCAATTTTAAAGAGAGAGGCTGTATTTCTGACCACCTCACTATATGGAG	4596
Db	4678	TTGCAACTGCAATTTTAAAGAGAGAGGCTGTATTTCTGACCACCTCACTATATGGAG	4737
QY	4597	AGGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGGCAAGTATGATTCG	4656
Db	4738	AGGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGGCAAGTATGATTCG	4797
QY	4657	GAAACGATCAACATCTAAGAGTAAATTTGAAAAAGGCTACTTTTTTGGAAATTAATTTGA	4716
Db	4798	GAAACGATCAACATCTAAGAGTAAATTTTGGAAAAAGGCTACTTTTTTGGAAATTAATTTGA	4857
QY	4717	AGGACTGGATAGAAAACTTAGAAGTAGACCGCTTTCAAGAGAGAAATTCAGTATATTTTCC	4776

Db	4858	AGGACTGATAGAAAACCTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTC		4917	
Qy	4777	CAATGCAAGCGCTCAGAAAAGTTTCTCTATTTTGGCTTAPAAATTCCTAAGGAAG		4836	
Db	4918	CAATGCAAGCGCTCAGAAAAGTTTCTCTATTTTGGCTTAPAAATTCCTAAGGAAG		4977	
Qy	4837	ATGTTCACTCCCTTTCAAAATCTTTTAAAGCTGGGAAGCTAAACATGCTTTTGCCA		4896	
Db	4978	ATGTTCACTCCCTTTCAAAATCTTTTAAAGCTGGGAAGCTAAACATGCTTTTGCCA		5037	
Qy	4897	TGGAAGATATAGCTTTTCTCAAGCAATTTGGAACAGGTTTGTAGAACTCACTAAAG		4956	
Db	5038	TGGAAGATATAGCTTTTCTCAAGCAATTTGGAACAGGTTTGTAGAACTCACTAAAG		5097	
Qy	4957	AACAGAGGAGAGATAATAGTTGTGGAACCTTTAAACAGACACACTTTGGTGGGAACGAA		5016	
Db	5098	AACAGAGGAGAGATAATAGTTGTGGAACCTTTAAACAGACACACTTTGGTGGGAACGAA		5157	
Qy	5017	CACAAGAGATAGAGTAGATATTTGAATTTGTATTTGTTCGGTCTGCTTACCTGGACTTCT		5076	
Db	5158	CACAAGAGATAGAGTAGATATTTGAATTTGTATTTGTTCGGTCTGCTTACCTGGACTTCT		5217	
Qy	5077	TTCTTTTCACTTAATTTTAACTTTGGTTTAAAGATTTT		5117	
Db	5218	TTCTTTTCACTTAATTTTAACTTTGGTTTAAAGATTTT		5258	
RESULT 7					
AX417824					
LOCUS					
DEFINITION					
Sequence 1 from Patent WO0231147.					
ACCESSION					
AX417824					
VERSION					
AX417824.1 GI:215222942					
KEYWORDS					
SOURCE					
ORGANISM					
Homo sapiens (human)					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
1					
Hu.Y. and Nepomichy, B.					
Human transporter proteins and polynucleotides encoding the same					
Patent: WO 0231147-A 1 18-APR-2002;					
LEXICON GENETICS INC (US)					
FEATURES					
source					
1. 4929					
/organism="Homo sapiens"					
/mol_type="unassigned DNA"					
/db_xref="taxon:9606"					
ORIGIN					
Query Match					
Best Local Similarity 99.9%; Score 4572; DB 6; Length 4929;					
Matches 4922; Conservative 0; Mismatches 7; Indels 0; Gaps 0;					
Qy	114	ATGTCCTGCAATTTAGGAGGTAGGAGTTGGAGACAGACAGAACACTTCTACTGAAG	173		
Db	1	ATGTCCTGCAATTTAGGAGGTAGGAGTTGGAGACAGACAGAACACTTCTACTGAAG	60		
Qy	174	AATTACTTAAATTAATGACAGAACCAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA	233		
Db	61	AATTACTTAAATTAATGACAGAACCAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA	120		
Qy	234	TTTTTTTTTATTTGGTTAATTAATTAATAGCATGATGATCCAAATAAGAAATATGAAGAA	293		
Db	121	TTTTTTTTTATTTGGTTAATTAATTAATAGCATGATGATCCAAATAAGAAATATGAAGAA	180		
Qy	294	GTGCCATATATAGAACTCAATCCCTATGACAAAGTTTACTCTTCTTAATCTAATTTCTTGA	353		
Db	181	GTGCCATATATAGAACTCAATCCCTATGACAAAGTTTACTCTTCTTAATCTAATTTCTTGA	240		
Qy	354	TATACTCCAGTGACTAATATTACAGCAGCATCATGAGAAAGTGCTACTGATCATCTA	413		

Db	241	TATACTCCAGTGACTAATATTACAGCAGCATCATGCAGAAAGTGCTACTGATCATCTA	300		
Qy	414	CTGATGTCATTAATTTACTGAGAAATATACAAATGAAAAGAAATGTTAACTCCAGTCTC	473		
Db	301	CTGATGTCATTAATTTACTGAGAAATATACAAATGAAAAGAAATGTTAACTCCAGTCTC	360		
Qy	474	TCCTAAGCCGAGCAACTTTGTAGGTGTGTTTTTCAAGAGACTCCATGCTCTATGAATTCGT	533		
Db	361	TCCTAAGCCGAGCAACTTTGTAGGTGTGTTTTTCAAGAGACTCCATGCTCTATGAATTCGT	420		
Qy	534	TTTTTTCTCGATATGATTCAGTATCTTCTATTTATATGATTTCAAGAGCTGGCTGTTC	593		
Db	421	TTTTTTCTCGATATGATTCAGTATCTTCTATTTATATGATTTCAAGAGCTGGCTGTTC	480		
Qy	594	AAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTTCAAGTTTTCAAGAGATCCCAT	653		
Db	481	AAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTTCAAGTTTTCAAGAGATCCCAT	540		
Qy	654	GATGTCGCCATTTATACAGTTGAAACCAATGTTCTCTTTGGAAGAGCTGGAGTCAACT	713		
Db	541	GATGTCGCCATTTATACAGTTGAAACCAATGTTCTCTTTGGAAGAGCTGGAGTCAACT	600		
Qy	714	AAAGCTGTTATTATGGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGGAGTA	773		
Db	601	AAAGCTGTTATTATGGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGGAGTA	660		
Qy	774	ATTTTAATATACCTAGTTATAGCATTTTCACTTTTGGATCTTTTGGCAATTCATATC	833		
Db	661	ATTTTAATATACCTAGTTATAGCATTTTCACTTTTGGATCTTTTGGCAATTCATATC	720		
Qy	834	GTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATGATACT	893		
Db	721	GTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATGATACT	780		
Qy	894	GCCTTTTGGCTTCTGGTCTCTATATACAGTTTAAATTTTCTTATGTCCTCTCT	953		
Db	781	GCCTTTTGGCTTCTGGTCTCTATATACAGTTTAAATTTTCTTATGTCCTCTCT	840		
Qy	954	ATGGCAGCATTTGCGACAGCTCTTTTGTATTTCTCTCAAGTAGCAGCATTTGTATATT	1013		
Db	841	ATGGCAGCATTTGCGACAGCTCTTTTGTATTTCTCTCAAGTAGCAGCATTTGTATATT	900		
Qy	1014	CTGCTTTTTTCTTTATGGAATTCATCTGTATTTTTTGTCTTAACTGACACCTCTT	1073		
Db	901	CTGCTTTTTTCTTTATGGAATTCATCTGTATTTTTTGTCTTAACTGACACCTCTT	960		
Qy	1074	TTTAAATAATCAAAACATGTGGATAGTGAATTTTGTCTGTCTGCTTTTGGATTT	1133		
Db	961	TTTAAATAATCAAAACATGTGGATAGTGAATTTTGTCTGTCTGCTTTTGGATTT	1020		
Qy	1134	ATTTGGCTTTATGATTAATCTCATAGAAAAGTTTTTCCAAATCGTTAGTGTGGCTTTTCAGT	1193		
Db	1021	ATTTGGCTTTATGATTAATCTCATAGAAAAGTTTTTCCAAATCGTTAGTGTGGCTTTTCAGT	1080		
Qy	1194	CTTTTCTGTCAGTACTTTTGTGATTTGATTTGACAGAGTCATGCAATTTAGAGATTTT	1253		
Db	1081	CTTTTCTGTCAGTACTTTTGTGATTTGATTTGACAGAGTCATGCAATTTAGAGATTTT	1140		
Qy	1254	AATGAAGTGTCTTCAATTTTCAAAATTTGACTGAGGCCCATATCTCTTAATTTATCAAAAT	1313		
Db	1141	AATGAAGTGTCTTCAATTTTCAAAATTTGACTGAGGCCCATATCTCTTAATTTATCAAAAT	1200		
Qy	1314	ATCTGCTCAGACTTAATAGTATATCTATGCTCTTGGCTGTCTATCTTGATCAAGTC	1373		
Db	1201	ATCTGCTCAGACTTAATAGTATATCTATGCTCTTGGCTGTCTATCTTGATCAAGTC	1260		
Qy	1374	ATTTCAAGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTGAAGCCCTTCATATTGG	1433		
Db	1261	ATTTCAAGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTGAAGCCCTTCATATTGG	1320		
Qy	1434	TCAAAGAGCAAAAGAAATTTAGAGAGTTATCAGAGGCAATGTTAATGGAATTTAGT	1493		
Db	1321	TCAAAGAGCAAAAGAAATTTATGAGAGTTATCAGAGGCAATGTTAATGGAATTTAGT	1380		

1494 TTTAGTGAAATTAATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAACCATAGAATT 1553  
1381 TTTAGTGAAATTAATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAACCATAGAATT 1440  
1554 AGTGGTATTCAGAAGACATACAGAAGAGGGTGAAATGTGGAGGCTTTGAGAAATTTG 1613  
1441 AGTGGTATTCAGAAGACATACAGAAGAGGGTGAAATGTGGAGGCTTTGAGAAATTTG 1500  
1614 TCATTTGACATATATAGGGGTGAGATTTACTGCTTTACTTTGGCCACAGTGGAAACAGGAAG 1673  
1501 TCATTTGACATATATAGGGGTGAGATTTACTGCTTTACTTTGGCCACAGTGGAAACAGGAAG 1560  
1674 AGTACATTTGATGAATTAATTTCTTTGGGACTCTGCCACCTCTCTGATGGGTTTGCATCTATA 1733  
1561 AGTACATTTGATGAATTAATTTCTTTGGGACTCTGCCACCTCTCTGATGGGTTTGCATCTATA 1620  
1734 TATGACACAGAGTCTCAGAAATAGATCAAAATGTTTGAAGCAAGAAATGATTTGGCAT 1793  
1621 TATGACACAGAGTCTCAGAAATAGATCAAAATGTTTGAAGCAAGAAATGATTTGGCAT 1580  
1794 TGTCCACAGTTAGATATACATTTGATGTTTGAAGTGTGAGAGAAATTTATCAATTTTG 1853  
1681 TGTCCACAGTTAGATATACATTTGATGTTTGAAGTGTGAGAGAAATTTATCAATTTTG 1740  
1854 GCTTCAATCAAGGGATACAGCCCAACATATATACAGAGTGTGAGAGTGTTTACTA 1913  
1741 GCTTCAATCAAGGGATACAGCCCAACATATATACAGAGTGTGAGAGTGTTTACTA 1800  
1914 GATTTAGACATGACAGTATCAAAAGATTAACCAAGCTAAATAATTAAGTGGTGTGCAAAAA 1973  
1801 GATTTAGACATGACAGTATCAAAAGATTAACCAAGCTAAATAATTAAGTGGTGTGCAAAAA 1860  
1974 AAGAGCTGTCTATAGGAATTTCTGTTCTGGGAACCCCAAGATACCTGCTAGATGAA 2033  
1861 AAGAGCTGTCTATAGGAATTTCTGTTCTGGGAACCCCAAGATACCTGCTAGATGAA 1920  
2034 CCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATAACAGA 2093  
1921 CCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATAACAGA 1980  
2094 AAGCCAAATCGGTGACAGTGTTCAGTACTCATTTTCATGAGTGAAGTGCATTTTGCA 2153  
1981 AAGCCAAATCGGTGACAGTGTTCAGTACTCATTTTCATGAGTGAAGTGCATTTTGCA 2040  
2154 GATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTTGTTGTTCTTCAATGTTCTCTC 2213  
2041 GATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTTGTTGTTCTTCAATGTTCTCTC 2100  
2214 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGTCACCA 2273  
2101 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGTCACCA 2160  
2274 GAATCTCTTTCTTCACTGTTTAAACACATATACCTGGAGTCTTTTATACACAGAAAT 2333  
2161 GAATCTCTTTCTTCACTGTTTAAACACATATACCTGGAGTCTTTTATACACAGAAAT 2220  
2334 GACCAACAACTGTGTATAGTTCCTTTTCAAGGACATGACAAATTTTCAAGTGTGTTT 2393  
2221 GACCAACAACTGTGTATAGTTCCTTTTCAAGGACATGACAAATTTTCAAGTGTGTTT 2280  
2394 TCTGCCCTPAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGGTGTTCCTATGACGACT 2453  
2281 TCTGCCCTPAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGGGTTTCCATGACGACT 2340  
2454 TTGGAAGACGATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTTAGT 2513  
2341 TTGGAAGACGATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTTAGT 2400  
2514 GTATTTACTACAGAGCCATGAGAGAAAGAAATGGAATTTCAAAATCTTTTGTATGAAATGGAA 2573  
2401 GTATTTACTACAGAGCCATGAGAGAAAGAAATGGAATTTCAAAATCTTTTGTATGAAATGGAA 2460

2574 CAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTTCTCTAGTGAGCACCACCATGAGCCTTTGG 2633  
2461 CAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTKCTCTAGTGAGCACCACCATGAGCCTTTGG 2520  
2634 AAACAACAGATGTATACAATAGCAAGTTTCATTTCTTTACCTTTGAAACGTTGAAAGTAA 2693  
2521 AAACAACAGATGTATACAATAGCAAGTTTCATTTCTTTACCTTTGAAACGTTGAAAGTAA 2580  
2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTT 2753  
2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTT 2640  
2754 TTGGTTCATCATCTTTTAAATAGTGTGGTTCCTCAATCAAACTGTTCACAGCTTATAT 2813  
2641 TTGGTTCATCATCTTTTAAATAGTGTGGTTCCTCAATCAAACTGTTCACAGCTTATAT 2700  
2814 TTTCTAAAACCTGGAGACAAACACACATAAATACAAAACAAAGCTGCTTCTTCAAAATCT 2873  
2701 TTTCTAAAACCTGGAGACAAACACACATAAATACAAAACAAAGCTGCTTCTTCAAAATCT 2760  
2874 GGTGACTCAGATATCAGTGTATTTAGCTTTTTCACAAAGCCAGAACATAATGTTGACG 2933  
2761 GGTGACTCAGATATCAGTGTATTTAGCTTTTTCACAAAGCCAGAACATAATGTTGACG 2820  
2934 ATGATTAATGACAGTGTATGTCGCTGCTCCCAAGTGTGCGCTTTTAAATGTTGATG 2993  
2821 ATGATTAATGACAGTGTATGTCGCTGCTCCCAAGTGTGCGCTTTTAAATGTTGATG 2880  
2994 CATTCAGAAAAGGACTATGTTTTGAGCTGTTTTTCAACAGTACTATGTTTATTTCTTTA 3053  
2881 CATTCAGAAAAGGACTATGTTTTGAGCTGTTTTTCAACAGTACTATGTTTATTTCTTTA 2940  
3054 CCTATATAGTGAATATCATTAAGTACTACTATCTTTATCATTTAAATGTTGACTGAAAC 3113  
2941 CCTATATAGTGAATATCATTAAGTACTACTATCTTTATCATTTAAATGTTGACTGAAAC 3000  
3114 ATCCAGATCTGAGTACCCCATTTTCAAGAAATTAAGTATAGTTTAAAAATTTGAG 3173  
3001 ATCCAGATCTGAGTACCCCATTTTCAAGAAATTAAGTATAGTTTAAAAATTTGAG 3060  
3174 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCAATGTTTACTGCAATGCCACTTACTTTGCC 3233  
3061 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCAATGTTTACTGCAATGCCACTTACTTTGCC 3120  
3234 ATGGAATATGACAGAAATCATAAGATCAAGCTTATATCTCACTTAAACTTTTCAAGTCTT 3293  
3121 ATGGAATATGACAGAAATCATAAGATCAAGCTTATATCTCACTTAAACTTTTCAAGTCTT 3180  
3294 TTGCCATCTGCATATTTGGAATGGACAAGCTGTTGTTGATATCCCTTATTTTATATCAT 3353  
3181 TTGCCATCTGCATATTTGGAATGGACAAGCTGTTGTTGATATCCCTTATTTTATATCAT 3240  
3354 CTATTTTGTATGCTAGGAGCTTATTTGGCAATTTCAATTTGGAATATATTTTATATCTGTA 3413  
3241 CTATTTTGTATGCTAGGAGCTTATTTGGCAATTTCAATTTGGAATATATTTTATATCTGTA 3300  
3414 AAGTTCCTTGTGCTGTTTTTGTGCTTATTTGCTTATGTTTCCATGATTTATCTGTTCACT 3473  
3301 AAGTTCCTTGTGCTGTTTTTGTGCTTATTTGCTTATTTGCTTATTTCCATGATTTCTGTTCACT 3360  
3474 TATATTTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3533  
3361 TATATTTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
3534 TATTTCTGTGGCAGCGTGTGCTTGTATTTGCAATCACTGAAATAAATCTTTTATGGGATAC 3593  
3421 TATTTCTGTGGCAGCGTGTGCTTGTATTTGCAATCACTGAAATAAATCTTTTATGGGATAC 3480  
3594 ACAATTTGCAATTTCTTCAATTTGCTTGTATCTCATCTCAATCTCAATCTCAATCTCA 3653  
3481 ACAATTTGCAATTTCTTCAATTTGCTTGTATCTCATCTCAATCTCAATCTCAATCTCA 3540  
3654 GGTGCTGCTGATTTCTTTCATAAAGATTTCTTGGAAAGATGTACGAAATAATGTTGACACC 3713



```
Db 3541 GGTGGCTGATTTCTTTTCATAGATTTCTTGGAGAAATGTACGAAATATGTGCACACC 3600
QY TATAATCCATCGGATAGGCTTTTCAGTAGCTGTATATCCGCTTACTGCGAGTGTGTACTG 3773
Db TATAATCCATCGGATAGGCTTTTCAGTAGCTGTATATCCGCTTACTGCGAGTGTGTACTG 3660
QY TGGATTTTCTCTTCAATACTACTAGAAAAAATATGGAGGCGAGATCAATAAGAAAAAGAT 3833
Db TGGATTTTCTCTTCAATACTACTAGAAAAAATATGGAGGCGAGATCAATAAGAAAAAGAT 3720
QY CCCTTTTTCAGAAACCTTTCAAGAAAGTCTAAAAATAGGAAGCTTCCAGAACCCACAGAC 3893
Db CCCTTTTTCAGAAACCTTTCAAGAAAGTCTAAAAATAGGAAGCTTCCAGAACCCACAGAC 3780
QY AATGAGGATCAAGATCAAGATGTCAAGCTGAAAGACTAAAGGTCAAGAGCTGTATGGGT 3953
Db AATGAGGATCAAGATGTCAAGCTGAAAGACTAAAGGTCAAGAGCTGTATGGGT 3840
QY TGCAGGTGTTGTGAGGAGAAACCATCCATTATGGTCAGCAATTTGTCATTAAGAAATATGAT 3900
Db TGCAGGTGTTGTGAGGAGAAACCATCCATTATGGTCAGCAATTTGTCATTAAGAAATATGAT 3840
QY GACAGAAAGATTTCTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 4073
Db GACAGAAAGATTTCTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 3960
QY TTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGGTCTGCGAAAAAGC 4133
Db TTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGGTCTGCGAAAAAGC 4020
QY ACAATTATTAATTTCTGTTGTGATATTAAGCAACTTTCAGGCCAGGTATTTTAGGA 4193
Db ACAATTATTAATTTCTGTTGTGATATTAAGCAACTTTCAGGCCAGGTATTTTAGGA 4080
QY GATTATTTCTTCAGACAGCAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTACTGTCT 4253
Db GATTATTTCTTCAGACAGCAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTACTGTCT 4140
QY CAGATAACCTTTGTGGCCAGATACATACATTGCGAGAACATTTGAAATTTATGGAGCT 4313
Db CAGATAACCTTTGTGGCCAGATACATACATTGCGAGAACATTTGAAATTTATGGAGCT 4200
QY GTCAAAGGAATGAGTGCAGTGCATGAAGAAAGTCAATAGTGCATGAATACACATGCATCT 4373
Db GTCAAAGGAATGAGTGCAGTGCATGAAGAAAGTCAATAGTGCATGAATACACATGCATCT 4260
QY GATTAAAAAGACATCTTCAGAAAGACTGTAAAGAAACTACCTGCGAGGAATCAAGCGAAG 4433
Db GATTAAAAAGACATCTTCAGAAAGACTGTAAAGAAACTACCTGCGAGGAATCAAGCGAAG 4320
QY TTGTGTTTCTCTTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCATCT 4493
Db TTGTGTTTCTCTTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCATCT 4380
QY ACAGGTATGATCCCAAGCCAAACAGACATGTGCGAGCAATTCGAACTGCAATTTAAA 4553
Db ACAGGTATGATCCCAAGCCAAACAGACATGTGCGAGCAATTCGAACTGCAATTTAAA 4440
QY AACAGAAAGCGGCTGTATTTCTGACCACTCATATATGAGGAGGCGAGGCTGTCTGT 4613
Db AACAGAAAGCGGCTGTATTTCTGACCACTCATATATGAGGAGGCGAGGCTGTCTGT 4500
QY GATCGAGTATCATATGATGTTGTGCGCAGTTAAGATGTATCGGAACAGTACAAATCTA 4673
Db GATCGAGTATCATATGATGTTGTGCGCAGTTAAGATGTATCGGAACAGTACAAATCTA 4560
QY AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTAAGAGCTGGTATGAAGAC 4733
Db AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTAAGAGCTGGTATGAAGAC 4620
QY CTAGAAGTAGACCGCCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 4793
```

```
Db 4621 CTAGAAGTAGACCGCCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 4680
QY GAAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCAAGTCCCTTCA 4853
Db GAAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCAAGTCCCTTCA 4740
QY CAATCTTTTTTAAAGCTGGAAGAAAGCTAAACATGCTTTTGGCAATGAAGAAATATAGCTTT 4913
Db CAATCTTTTTTAAAGCTGGAAGAAAGCTAAACATGCTTTTGGCAATGAAGAAATATAGCTTT 4800
QY TCTCAACACATGGAAGAAAGCTTTTGTAGAACTCACTAAAGCAAGAGGAGGAGAT 4973
Db TCTCAACACATGGAAGAAAGCTTTTGTAGAACTCACTAAAGCAAGAGGAGGAGAT 4860
QY AATAGTTGTGGAACCTTTTAAACAGCACACTTTTGTGGGAAACGAAACACAAAGAGATAGTA 5033
Db AATAGTTGTGGAACCTTTTAAACAGCACACTTTTGTGGGAAACGAAACACAAAGAGATAGTA 4920
QY 5034 GTATTTTGA 5042
Db 4921 GTATTTTGA 4929

RESULT 8
AX417826
LOCUS Sequence 3 from Patent WO0231147. 4785 bp DNA linear PAT 18-JUN-2002
DEFINITION AX417826
ACCESSION AX417826
VERSION AX417826.1 GI:21522943
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1
Hu, Y. and Nepomnichy, B.
Human transporter proteins and polynucleotides encoding the same
Patent: WO 0231147-A 3 18-APR-2002;
LEXICON GENETICS INC (US)

FEATURES
source
1..4785
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 80.5%; Score 4408; DB 6; Length 4785;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 4758; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 114 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCAGAACACTTCTACTGAAG 173
Db 1 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCAGAACACTTCTACTGAAG 60
QY 174 AATTACTTAATTAATTCAGAACCCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 233
Db 61 AATTACTTAATTAATTCAGAACCCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 120
QY 234 TTTTCTTTTATTTTGGTTAAATTAATTAAATTAGCATGATGCATCCAAATTAAGAAATATGAAGA 293
Db 121 TTTTCTTTTATTTTGGTTAAATTAATTAAATTAGCATGATGCATCCAAATTAAGAAATATGAAGA 180
QY 294 GTGCGCTAATATAGAACTCAATCTCTATGCAAGTTTACTCTTTCTTATCTTAATTTCTTGGGA 353
Db 181 GTGCGCTAATATAGAACTCAATCTCTATGCAAGTTTACTCTTTCTTATCTTAATTTCTTGGGA 240
QY 354 TATCTCTCAGTGTCTAATATTTACAAAGCAGCATCATGCAAGAAAGTGTCTACTGATCATCTA 413
Db 241 TATCTCTCAGTGTCTAATATTTACAAAGCAGCATCATGCAAGAAAGTGTCTACTGATCATCTA 300
QY 414 CCTGATGTCAATATCTGGAAGAAATATACAAATGAAGAAAGATTTTAACATCCAGTCTC 473
Db 301 CCTGATGTCAATATCTGGAAGAAATATACAAATGAAGAAAGATTTTAACATCCAGTCTC 360
```





2521	Db	AAACAACAGATGTATACAATAGCAAAAGTTTCATTTCTTTTACCTTTGAAACGTGAAGATPAAA	2580
2694	Qy	TCAGTGAGATCAGTGTTCCTCTCTCTTTAAATTTTTTTTACAGTTCAGATTTTTTATGTTTT	2753
2581	Db	TCAGTGAGATCAGTGTTCCTCTCTCTTTAAATTTTTTTTACAGTTCAGATTTTTTATGTTTT	2640
2754	Qy	TTGGTTCATCACATCTTTTAAAAATGCTGGTCCCATCAAACTGTTTCCAGACTTATAT	2813
2641	Db	TTGGTTCATCACATCTTTTAAAAATGCTGGTCCCATCAAACTGTTTCCAGACTTATAT	2700
2814	Qy	TTTCTAAAACCTGGAGACAAACACATAAATACAAAAAACAAGTCGTCTTCTTCAAAATTCCT	2873
2701	Db	TTTCTAAAACCTGGAGACAAACACATAAATACAAAAAACAAGTCGTCTTCTTCAAAATTCCT	2760
2874	Qy	GCTGACTCAGATATCAGTGTATCTTATTAGCTTTTTTACAAGCCAGACACATAAATGGTGACG	2933
2761	Db	GCTGACTCAGATATCAGTGTATCTTATTAGCTTTTTTACAAGCCAGACACATAAATGGTGACG	2820
2934	Qy	ATGATTAAATGACAGTGACTATGTATCCGTGGCTCCCATATAGTGGGCTTTAAATGTGATG	2993
2821	Db	ATGATTAAATGACAGTGACTATGTATCCGTGGCTCCCATATAGTGGGCTTTAAATGTGATG	2880
2994	Qy	CATTTCAGAAAGGACTATGTTTTTGGCAGCTGTTTTTCAACAGTACTATGGTTTTATTCTTTA	3053
2881	Db	CATTTCAGAAAGGACTATGTTTTTGGCAGCTGTTTTTCAACAGTACTATGGTTTTATTCTTTA	2940
3054	Qy	CCATATTAGTGAATATCATTTAGTAACTACTATCTTTTATCATTTAAATGTGACTGAAACC	3113
2941	Db	CCATATTAGTGAATATCATTTAGTAACTACTATCTTTTATCATTTAAATGTGACTGAAACC	3000
3114	Qy	ATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCATATAGTATTTTTTAAAAATTGAG	3173
3001	Db	ATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCATATAGTATTTTTTAAAAATTGAG	3060
3174	Qy	CTGTATTTTCAAGCAGCTTTGCTTTGGAAATCATTTGTTTACTGCAATGCCACTTACTTTGCC	3233
3061	Db	CTGTATTTTCAAGCAGCTTTGCTTTGGAAATCATTTGTTTACTGCAATGCCACTTACTTTGCC	3120
3234	Qy	ATGGAAAATGCAGAGATCATTAAGATCAAAAGCTTTATATCTCAACTTAACTTTTCAAGTCTT	3293
3121	Db	ATGGAAAATGCAGAGATCATTAAGATCAAAAGCTTTATATCTCAACTTAACTTTTCAAGTCTT	3180
3294	Qy	TTGCCATCTGCATATTTGGATTTGGACACAGCTTTGTTTGATATGCCCTTATTTTTTATCATTT	3353
3181	Db	TTGCCATCTGCATATTTGGATTTGGACACAGCTTTGTTTGATATGCCCTTATTTTTTATCATTT	3240
3354	Qy	CTTATTTTGTATGCTPAGGAAGCTTATTCGGCATTTTCATTTATGATATATTTTTTATCTGTA	3413
3241	Db	CTTATTTTGTATGCTPAGGAAGCTTATTCGGCATTTTCATTTATGATATATTTTTTATCTGTA	3300
3414	Qy	AAGTTCCTTCGTGTGGTTTTTTTGGCTTATTTGGTTATGTTTCCATCAGTTATTTCTGTTCACT	3473
3301	Db	AAGTTCCTTCGTGTGGTTTTTTTGGCTTATTTGGTTATGTTTCCATCAGTTATTTCTGTTCACT	3360
3474	Qy	TATATTGCTTCTTTTCACTTTTAAAGAAAATTTTAAATACCAAAGAAATTTTGGTCAATTTATC	3533
3361	Db	TATATTGCTTCTTTTCACTTTTAAAGAAAATTTTAAATACCAAAGAAATTTTGGTCAATTTATC	3420
3534	Qy	TATTTCTGTGGCAGGTGGCTTGTATTTGCAATCATCTGAAATAACTTTCTTTTATGGGATAC	3593
3421	Db	TATTTCTGTGGCAGGTGGCTTGTATTTGCAATCATCTGAAATAACTTTCTTTTATGGGATAC	3480
3594	Qy	ACAAATGCAACTATCTTTTCATTAATGCGCTTTTGTATCATCTTCAAACTCATCCACTTCTA	3653
3481	Db	ACAAATGCAACTATCTTTTCATTAATGCGCTTTTGTATCATCTTCAAACTCATCCACTTCTA	3540
3654	Qy	GGTTGCCCTGATTTCTTTTCATAAAGATTTCTTTGGNAGAAATGTACGAAAAAATGTGGACACC	3713
3541	Db	GGTTGCCCTGATTTCTTTTCATAAAGATTTCTTTGGNAGAAATGTACGAAAAAATGTGGACACC	3600
3714	Qy	TATTAATCCATGGGATAGGCTTTTCAGTAGCTGTATTATTCGCGCTTACCTTCGAGTGTGTACTG	3773



Db 1681 TGATGAATATCTTTGTGACTCTGCCACCTTCTGATGGTTCATCATATATGAC 1740  
QY 1741 ACAGAGTCTCAGAAATAGATGAATGTTGAAGCAAGAAATAGATGGCATTTGTTCAC 1800  
Db 1741 ACAGAGTCTCAGAAATAGATGAATGTTGAAGCAAGAAATAGATGGCATTTGTTCAC 1800  
QY 1801 AGTTAGATATACATTTGATGTTTGGACATGAGAAATTTATCAATTTTGGCTTCAA 1860  
Db 1801 AGTTAGATATACATTTGATGTTTGGACATGAGAAATTTATCAATTTTGGCTTCAA 1860  
QY 1861 TCAAGGGATACAGCCAAATATAATCAAGAAAGTGCAGAGGTTTACTAGATTAG 1920  
Db 1861 TCAAGGGATACAGCCAAATATAATCAAGAAAGTGCAGAGGTTTACTAGATTAG 1920  
QY 1921 ACATGAGACTATCAAGATACCAAGCTTAAATTAAGTGGTGCCTCAAAAAGAGC 1980  
Db 1921 ACATGAGACTATCAAGATACCAAGCTTAAATTAAGTGGTGCCTCAAAAAGAGC 1980  
QY 1981 TGTCAATAGGAATGCTGTTGGGACCCCAAGATAGTCTGCTAGATGAACCAACAG 2040  
Db 1981 TGTCAATAGGAATGCTGTTGGGACCCCAAGATAGTCTGCTAGATGAACCAACAG 2040  
QY 2041 CTGGGAATGACCCCTGTTCTGCACATATGATGAATCTTTAAATACAGAAAGCCA 2100  
Db 2041 CTGGGAATGACCCCTGTTCTGCACATATGATGAATCTTTAAATACAGAAAGCCA 2100  
QY 2101 ATCGGGTGACAGTGTTCAGTACTCATTTATGAGATGAAGCTGACATTTCTGCAGATGA 2160  
Db 2101 ATCGGGTGACAGTGTTCAGTACTCATTTATGAGATGAAGCTGACATTTCTGCAGATGA 2160  
QY 2161 AGCTGTGATATACAGGAATGCTGAAATGTTGGTCTTCAATGTTCTCAAAAGTA 2220  
Db 2161 AGCTGTGATATACAGGAATGCTGAAATGTTGGTCTTCAATGTTCTCAAAAGTA 2220  
QY 2221 AATGGGGATCGGCTACCGCTGAGCATGTATACAGAAATATTTGTCACAGAAATCTC 2280  
Db 2221 AATGGGGATCGGCTACCGCTGAGCATGTATACAGAAATATTTGTCACAGAAATCTC 2280  
QY 2281 TTTCTTCACTGTTTAAACAATATACCTGAGCTACTTTATATCAACAGAAATGACCAAC 2340  
Db 2281 TTTCTTCACTGTTTAAACAATATACCTGAGCTACTTTATATCAACAGAAATGACCAAC 2340  
QY 2341 AACTTGTGTATAGTGTGCTTTCAAGGACATGACAAATTTTCAGTGTGTTTCTGCC 2400  
Db 2341 AACTTGTGTATAGTGTGCTTTCAAGGACATGACAAATTTTCAGTGTGTTTCTGCC 2400  
QY 2401 TAGACAGTCATTCAAATTTGGGTGCTATTTCTTATGGTGTTCATGACGACTTTGGAAG 2460  
Db 2401 TAGACAGTCATTCAAATTTGGGTGCTATTTCTTATGGTGTTCATGACGACTTTGGAAG 2460  
QY 2461 AGGTATTTTAAAGCTAGAGTTGAAGCAGAAATGACCAAGCATTTATAGTGTATTTA 2520  
Db 2461 AGGTATTTTAAAGCTAGAGTTGAAGCAGAAATGACCAAGCATTTATAGTGTATTTA 2520  
QY 2521 CTCAGAGCCACTGGAGGAAGAAATGGAATCCTTTTGAATGAATGAACAGAGCT 2580  
Db 2521 CTCAGAGCCACTGGAGGAAGAAATGGAATCCTTTTGAATGAATGAACAGAGCT 2580  
QY 2581 TACTTATTTCTGAAACCAAGGCTCTCTAGTGAGCACCAGCTTTGGAAACAC 2640  
Db 2581 TACTTATTTCTGAAACCAAGGCTCTCTAGTGAGCACCAGCTTTGGAAACAC 2640  
QY 2641 AGATGATACATAGCAAAAGTTTCAATTTTACCTTGAACCGTGAAGTAAATCAGTGA 2700  
Db 2641 AGATGATACATAGCAAAAGTTTCAATTTTACCTTGAACCGTGAAGTAAATCAGTGA 2700  
QY 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAATTTTATGTTTGGTTC 2760  
Db 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAATTTTATGTTTGGTTC 2760  
QY 2761 ATCACTCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTTCAGACTTATATTTCTAA 2820

Db 2761 ATCACTCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTTCAGACTTATATTTCTAA 2820  
QY 2821 AACCTGGAGCAAAACCAATATAATCAAAACAGTCTGCTTCTTCAAAATTTCTGCTG 2877  
Db 2821 AACCTGGAGCAAAACCAATATAATCAAAACAGTCTGCTTCTTCAAAATTTCTGCTG 2877  
RESULT 10  
LOCUS AX714300 3268 bp DNA linear PAT 15-APR-2003  
DEFINITION Sequence 984 from Patent EP1293569.  
ACCESSION AX714300  
VERSION AX714300.1 GI:29889252  
KEYWORDS Homo sapiens (human)  
SOURCE ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,  
Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,  
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and  
Masuho, Y.  
Full-length cDNA  
Patent: EP 1293569-A 984 19-MAR-2003;  
Helix Research Institute (JP); Research Association for  
Biotechnology (JP)  
FEATURES  
source  
1..3268  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 50.0%; Score 2739; DB 6; Length 3268;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3039; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 2386 GTTTGTTTCTGCCCTAGACAGTCATTCAAAATTTGGGTGTCATTTCTATGCTGTTCCA 2445  
Db 220 GTTTGTTTCTGCCCTAGACAGTCATTCAAAATTTGGGTGTCATTTCTATGCTGTTCCA 279  
QY 2446 TGACGACTTTGGAAGACGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATGACCAAGCAG 2505  
Db 280 TGACGACTTTGGAAGACGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATGACCAAGCAG 339  
QY 2506 ATTATAGTGTATTACTCAGCAGCCACTGGAGGAAGAAATGGAATCAAAATCTTTGATG 2565  
Db 340 ATTATAGTGTATTACTCAGCAGCCACTGGAGGAAGAAATGGAATCAAAATCTTTGATG 399  
QY 2566 AAATGAAACAGAGCTTACTTATTCTTCTGAAACCAAGGCTTCTAGTGAGCACCATGA 2625  
Db 400 AAATGAAACAGAGCTTACTTATTCTTCTGAAACCAAGGCTTCTAGTGAGCACCATGA 459  
QY 2626 GCCTTTGGAAACACAGATGTATACAAATAGAAAGTTTCTTACCTTGAACCGTG 2685  
Db 460 GCCTTTGGAAACACAGATGTATACAAATAGAAAGTTTCTTACCTTGAACCGTG 519  
QY 2686 AAAGTAAATCAGTGAGTCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAAGATT 2745  
Db 520 AAAGTAAATCAGTGAGTCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAAGATT 579  
QY 2746 TTATGTTTTTGGTTCATCATCTCTTTTAAATAGCTGTGGTTCCTCATCAAACTGTTCCAG 2805  
Db 580 TTATGTTTTTGGTTCATCATCTCTTTTAAATAGCTGTGGTTCCTCATCAAACTGTTCCAG 639  
QY 2806 ACTTATATTTTCTTAAACCTGGAGCAACACATATAAATACAAAACAGTCTGCTTCTTC 2865  
Db 640 ACTTATATTTTCTTAAACCTGGAGCAACACATATAAATACAAAACAGTCTGCTTCTTC 699  
QY 2866 AAAATTTCTGCTCAGATATCAGTGATCTTATTTAGCTTTTTCACAGCCAGAACATAA 2925  
Db 700 AAAATTTCTGCTCAGATATCAGTGATCTTATTTAGCTTTTTCACAGCCAGAACATAA 759

QY 2926 TGGTGACGATGATTAATGACAGTACTATGTATCGTGGCTCCCATAGTGGCGCTTAA 2985  
 Db 760 TGGTGACGATGATTAATGACAGTACTATGTATCGTGGCTCCCATAGTGGCGCTTAA 819  
 QY 2986 ATGTGATGATTCAGAAAAGGACTATGTTTGGAGCTGTTTCAACAGTACTAGGTTT 3045  
 Db 820 ATGTGATGATTCAGAAAAGGACTATGTTTGGAGCTGTTTCAACAGTACTAGGTTT 879  
 QY 3046 ATTCTTTTACCTATATTAGTGAATATCATTAGTAACTACTATCTTTATCATTTAAATGTA 3105  
 Db 880 ATTCTTTTACCTATATTAGTGAATATCATTAGTAACTACTATCTTTATCATTTAAATGTA 939  
 QY 3106 CTGAAACCATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTAAGTATAGTTTTTA 3165  
 Db 940 CTGAAACCATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTAAGTATAGTTTTTA 999  
 QY 3166 AAAATTGAGCTGATTTTCCAGCAGCTTTGCTTGGAACTATTGTTACTGCAATGCCACTT 3225  
 Db 1000 AAAATTGAGCTGATTTTCCAGCAGCTTTGCTTGGAACTATTGTTACTGCAATGCCACTT 1059  
 QY 3226 ACTTTGCCATGGAAATCGAGAAATCATAAGATCAAAAGCTTATACCTAAACTTTT 3285  
 Db 1060 ACTTTGCCATGGAAATCGAGAAATCATAAGATCAAAAGCTTATACCTAAACTTTT 1119  
 QY 3286 CAGGCTTTTGGCATCTGCATATTGGATTGGACAGCTGTTGTTGATATCCCTTATTTT 3345  
 Db 1120 CAGGCTTTTGGCATCTGCATATTGGATTGGACAGCTGTTGTTGATATCCCTTATTTT 1179  
 QY 3346 TTATCATTTCTTATTTTGTAGTGTAGGAAGCTTATTGGCACTTTCATTATGGATTATATTTT 3405  
 Db 1180 TTATCATTTCTTATTTTGTAGTGTAGGAAGCTTATTGGCACTTTCATTATGGATTATATTTT 1239  
 QY 3406 ATACTGTAAGTTCCTGCTGGTGGTTTGGCTTTATGTTGTTATGTTCCATCAGTTATTC 3465  
 Db 1240 ATACTGTAAGTTCCTGCTGGTGGTTTGGCTTTATGTTGTTATGTTCCATCAGTTATTC 1299  
 QY 3466 TGTTTCACTTATATTGCTTTTCACTTTAAGAAAAATTTAAATACCAAGAAATTTTGGT 3525  
 Db 1300 TGTTTCACTTATATTGCTTTTCACTTTAAGAAAAATTTAAATACCAAGAAATTTTGGT 1359  
 QY 3526 CATTTATCTATCTGTCGACAGGTTGGCTTGTATGTAATCAATCAGTAAATACCTTCTTTA 3585  
 Db 1360 CATTTATCTATCTGTCGACAGGTTGGCTTGTATGTAATCAATCAGTAAATACCTTCTTTA 1419  
 QY 3586 TGGGATACAAATGCAACTATCTTCAATTAATGCTTTTGTATCATCATCTCCATCTATC 3645  
 Db 1420 TGGGATACAAATGCAACTATCTTCAATTAATGCTTTTGTATCATCATCTCCATCTATC 1479  
 QY 3646 CACTTCTAGGTTGCCGTGATTTCTTTCAATAAGATTTCTTGGAGAATGTACGAAAAATG 3705  
 Db 1480 CACTTCTAGGTTGCCGTGATTTCTTTCAATAAGATTTCTTGGAGAATGTACGAAAAATG 1539  
 QY 3706 TGGACACCTATAATCCATGGGATAGGCTTTTCAGTAGCTGTTTATATCGCTTACCTGCAGT 3765  
 Db 1540 TGGACACCTATAATCCATGGGATAGGCTTTTCAGTAGCTGTTTATATCGCTTACCTGCAGT 1599  
 QY 3766 GTGTACTGTGGATTTTCTCTTCAATACTATGAGAAAAATATGGAGGAGATCAATAA 3825  
 Db 1600 GTGTACTGTGGATTTTCTCTTCAATACTATGAGAAAAATATGGAGGAGATCAATAA 1659  
 QY 3826 GAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGCTTAAATAATAGGAAGCTTCCGAAC 3885  
 Db 1660 GAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGCTTAAATAATAGGAAGCTTCCGAAC 1719  
 QY 3886 CACCAGACAATGAGGATCAAGATGAAGATGTCAAAGCTGAAAGCTAAAGGTCAAGAGC 3945  
 Db 1720 CACCAGACAATGAGGATCAAGATGAAGATGTCAAAGCTGAAAGCTAAAGGTCAAGAGC 1779  
 QY 3946 TGATGGGTTGCCAGTGTGTGAGGAGAAACCATCTATGTTGTCAGCAATTTGTCATTAAG 4005  
 Db 1780 TGATGGGTTGCCAGTGTGTGAGGAGAAACCATCTATGTTGTCAGCAATTTGTCATTAAG 1839

QY 4006 AATATGATGACAAAGAAAGATTTTCTTTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAAT 4065  
 Db 1840 AATATGATGACAAAGAAAGATTTTCTTTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAAT 1899  
 QY 4066 ACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATATGGGTCCAAATGTGCTG 4125  
 Db 1900 ACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATATGGGTCCAAATGTGCTG 1959  
 QY 4126 GCAAAAGCACAATTTAATATTTCTGTTGGTGATATTGAACCAACTTCAGGCGCAGTAT 4185  
 Db 1960 GCAAAAGCACAATTTAATATTTCTGTTGGTGATATTGAACCAACTTCAGGCGCAGTAT 2019  
 QY 4186 TTTTAGAGATATTTCTTTCAGAGACAAGTGAAGATGATGATTCACCTGAAGTGTATGGGTT 4245  
 Db 2020 TTTTAGAGATATTTCTTTCAGAGACAAGTGAAGTGAAGTGAAGTGAAGTGTATGGGTT 2079  
 QY 4246 ACTGTCTCTCAGATAAACCCTTTGTGGCCAGATATACATTGCGAGGAACATTTTGAATTT 4305  
 Db 2080 ACTGTCTCTCAGATAAACCCTTTGTGGCCAGATATACATTGCGAGGAACATTTTGAATTT 2139  
 QY 4306 ATGGAGCTGTCAAAGGAATGAGTGCAGATGACATGAAGAAGTCAATGAATTAACAC 4365  
 Db 2140 ATGGAGCTGTCAAAGGAATGAGTGCAGATGACATGAAGAAGTCAATGAATTAACAC 2199  
 QY 4366 ATGCACCTTGATTTTAAAGAACATCTTCAGAGACTGTAAAGAACTACCTGCGAGGAATCA 4425  
 Db 2200 ATGCACCTTGATTTTAAAGAACATCTTCAGAGACTGTAAAGAACTACCTGCGAGGAATCA 2259  
 QY 4426 AACGAAAGTTGTTTGTCTTAAGTATGCTTAGGGAATCCTCAGATTACTTTCTAGATG 4485  
 Db 2260 AACGAAAGTTGTTTGTCTTAAGTATGCTTAGGGAATCCTCAGATTACTTTCTAGATG 2319  
 QY 4486 AACCATCTACAGGTATGGATCCCAAGCCAAACAGACACATGTGGGAGCAATTCGAAGT 4545  
 Db 2320 AACCATCTACAGGTATGGATCCCAAGCCAAACAGACACATGTGGGAGCAATTCGAAGT 2379  
 QY 4546 CATTTAAAAACAGAAAGCGGCTGTATTCTGACCACTCACTATATGAGAGGAGGAGAG 4605  
 Db 2380 CATTTAAAAACAGAAAGCGGCTGTATTCTGACCACTCACTATATGAGAGGAGGAGAG 2439  
 QY 4606 CTGTCTGTGATCGAGTAGCTATCATGTGCTGTGGGAGTTAAGATGTATCGGAAACAGTAC 4665  
 Db 2440 CTGTCTGTGATCGAGTAGCTATCATGTGCTGTGGGAGTTAAGATGTATCGGAAACAGTAC 2499  
 QY 4666 AACATCTAAAGAGTAAATTTGAAAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGA 4725  
 Db 2500 AACATCTAAAGAGTAAATTTGAAAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGA 2559  
 QY 4726 TAGAAAACTTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCAA 4785  
 Db 2560 TAGAAAACTTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCAA 2619  
 QY 4786 GCGTCAGGAAAGTTTCTTCTTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCACT 4845  
 Db 2620 GCGTCAGGAAAGTTTCTTCTTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTCACT 2679  
 QY 4846 CCCTTTCACAATCTTTTAAAGCTGGAAGAAGCTTAAACATGCTTTTGCCATTTGAAGAT 4905  
 Db 2680 CCCTTTCACAATCTTTTAAAGCTGGAAGAAGCTTAAACATGCTTTTGCCATTTGAAGAT 2739  
 QY 4906 ATAGCTTTTCTCAAGCAATTTGGAACAGGTTTTTGTAGAACTCACTAAGAACAGAGG 4965  
 Db 2740 ATAGCTTTTCTCAAGCAATTTGGAACAGGTTTTTGTAGAACTCACTAAGAACAGAGG 2799  
 QY 4966 AGGAAGATTAATAGTTGTGAACTTTTAAACAGCACACTTTTGGTGGGAAACGAAACAGAG 5025  
 Db 2800 AGGAAGATTAATAGTTGTGAACTTTTAAACAGCACACTTTTGGTGGGAAACGAAACAGAG 2859  
 QY 5026 ATAGAGTAGTATTTGAATTTGTTATGTTGCTGCTTACTGCGGACTTCTTTCTTTTC 5085  
 Db 2860 ATAGAGTAGTATTTGAATTTGTTATGTTGCTGCTTACTGCGGACTTCTTTCTTTTC 2919  
 QY 5086 ACTTAATTTTAACTTTGGTTTTAAAAAGTTTTTTTATTTGGAAATGTTAACTGGAGAACCAAGA 5145



Db 880 ATTCCTTACCTATATAGTGAATCATATAGTAACACTACTATCTTTATCATTTAAATGGA 939  
Qy 3106 CTGAAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTAAGTATAGTTTAA 3165  
Db 940 CTGAAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTAAGTATAGTTTAA 999  
Qy 3166 AAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTCAATGCCACCTT 3225  
Db 1000 AAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTCAATGCCACCTT 1059  
Qy 3226 ACTTTGCCATGGAATCCAGAGAAATCATTAAGATCAAAAGCTTTATCTCAAACTTAACCTTT 3285  
Db 1060 ACTTTGCCATGGAATCCAGAGAAATCATTAAGATCAAAAGCTTTATCTCAAACTTAACCTTT 1119  
Qy 3286 CAGGCTTTTGGCATCTGCATATTTGGATTGGACAGCTTTGTTGATATCCCTTATTTT 3345  
Db 1120 CAGGCTTTTGGCATCTGCATATTTGGATTGGACAGCTTTGTTGATATCCCTTATTTT 1179  
Qy 3346 TTATCATTTCTTATTTTGATGCTAGGAAGCTTTATTTGGCATTTTCATTTATGGATTAATTTTT 3405  
Db 1180 TTATCATTTCTTATTTTGATGCTAGGAAGCTTTATTTGGCATTTTCATTTATGGATTAATTTTT 1239  
Qy 3406 ATACTGTAAAGTCTTCTGCTGGTGTGTTTGGCTTATTTGGTTATGTTCCATCAGTTATTC 3465  
Db 1240 ATACTGTAAAGTCTTCTGCTGGTGTGTTTGGCTTATTTGGTTATGTTCCATCAGTTATTC 1299  
Qy 3466 TGTTCACCTTATATGCTTCTTTCACTTTTAAGAAATTTAAATACCAAGAAATTTTGGT 3525  
Db 1300 TGTTCACCTTATATGCTTCTTTCACTTTTAAGAAATTTAAATACCAAGAAATTTTGGT 1359  
Qy 3526 CATTTATCTATTTCTGTGGCAGCTTGGCTTGTATTTGCAATCACTGAATAAATCTTTCTTTA 3585  
Db 1360 CATTTATCTATTTCTGTGGCAGCTTGGCTTGTATTTGCAATCACTGAATAAATCTTTCTTTA 1419  
Qy 3586 TGGGATACAAATTCGAATTTCTTCAATATGCTTTTGTATGCTTATCATCTTCCATCTATC 3645  
Db 1420 TGGGATACAAATTCGAATTTCTTCAATATGCTTTTGTATGCTTATCATCTTCCATCTATC 1479  
Qy 3646 CACTTCTAGGTTGCTGTATTTCTTCAATAAGATTTCTTGAAGAATGTACGAAAAAATG 3705  
Db 1480 CACTTCTAGGTTGCTGTATTTCTTCAATAAGATTTCTTGAAGAATGTACGAAAAAATG 1539  
Qy 3706 TGGACACCTATAATCCATGGGATAGCTTTTCAGTAGCTGTTATATCGCTTTACCTGCAGT 3765  
Db 1540 TGGACACCTATAATCCATGGGATAGCTTTTCAGTAGCTGTTATATCGCTTTACCTGCAGT 1599  
Qy 3766 GTGTACTGTGGATTTTCTCTTACAAATCTATAGAAAAAATATGGAGGCGAGATCAATAA 3825  
Db 1600 GTGTACTGTGGATTTTCTCTTACAAATCTATAGAAAAAATATGGAGGCGAGATCAATAA 1659  
Qy 3826 GAAAGATCCCTTTTTCAGAACCTTTTCAAGAGCTTCAAAAATAGGAAGCTTCCAGAAC 3885  
Db 1660 GAAAGATCCCTTTTTCAGAACCTTTTCAAGAGCTTCAAAAATAGGAAGCTTCCAGAAC 1719  
Qy 3886 CACCAGACAATGAGGATGAAGTAAAGATGTCAAAAGCTTGAAGACTAAAGCTCAAGAGC 3945  
Db 1720 CACCAGACAATGAGGATGAAGTAAAGATGTCAAAAGCTTGAAGACTAAAGCTCAAGAGC 1779  
Qy 3946 TGAATGGTTGCCAGTTGTTGAGGAGAAACCAATTCATTTATGGTCAGCAATTTGCATTAAG 4005  
Db 1780 TGAATGGTTGCCAGTTGTTGAGGAGAAACCAATTCATTTATGGTCAGCAATTTGCATTAAG 1839  
Qy 4006 AATATGATGACAAAGAAATTTCTTCTTTCAAGAAAAATGAAAGAGTGGCACTAAAT 4065  
Db 1840 AATATGATGACAAAGAAATTTCTTCTTTCAAGAAAAATGAAAGAGTGGCACTAAAT 1899  
Qy 4066 ACATCTCTTTCTGTGTAAAAAGAGAGATCTTAGGCTATTTGGGTCCAAATGGTGTCTG 4125  
Db 1900 ACATCTCTTTCTGTGTAAAAAGAGAGATCTTAGGCTATTTGGGTCCAAATGGTGTCTG 1959  
Qy 4126 GCAAAAGCACAATTAATTAATTTCTGGTTGGTATTTGAACCACTTTCAGGCAGGATAT 4185

Db 1960 GCAAAAGCACAATTAATTAATTTCTGGTTGGTGATATTGAACCAACTTTCAGGCCAGGTAT 2019  
Qy 4186 TTTTAGAGAGATTAATTTCTTCAGAGACAAGTGAAGATGATGATTCACTGAAGTGATGGGTT 4245  
Db 2020 TTTTAGAGAGATTAATTTCTTCAGAGACAAGTGAAGATGATGATTCACTGAAGTGATGGGTT 2079  
Qy 4246 ACTGTCTCTCAGATAAAACCCCTTTGTGCCAGATACTACATTTCAGGGAACATTTTGAATTT 4305  
Db 2080 ACTGTCTCTCAGATAAAACCCCTTTGTGCCAGATACTACATTTCAGGGAACATTTTGAATTT 2139  
Qy 4306 ATGGAGCTGTCAAAAGGAATGAGTGAAGTGAACATGAAGAAAGTCAATGTCGAATTAACAC 4365  
Db 2140 ATGGAGCTGTCAAAAGGAATGAGTGAAGTGAACATGAAGAAAGTCAATGTCGAATTAACAC 2199  
Qy 4366 ATGCACCTTGATTTTAAAGAACATCTTTCAAGAGACTGTAAAGAACTACCTGCAGGAATCA 4425  
Db 2200 ATGCACCTTGATTTTAAAGAACATCTTTCAAGAGACTGTAAAGAACTACCTGCAGGAATCA 2259  
Qy 4426 AACGAAAGTTGTGTTTTGCTCTAAGTATGCTAGGGAATCCTCAGATTAATCTTTGCTAGATG 4485  
Db 2260 AACGAAAGTTGTGTTTTGCTCTAAGTATGCTAGGGAATCCTCAGATTAATCTTTGCTAGATG 2319  
Qy 4486 ARCCATCTACAGTATGGATCCCARAGCCAAAACAGCACATGTGGCAGCAATTTTCCGAACCTG 4545  
Db 2320 ARCCATCTACAGTATGGATCCCARAGCCAAAACAGCACATGTGGCAGCAATTTTCCGAACCTG 2379  
Qy 4546 CATTTAAAAACAGAAAGCGGCTGTATTTCTGACCACTCACTATATGGAGGAGCGAGAG 4605  
Db 2380 CATTTAAAAACAGAAAGCGGCTGTATTTCTGACCACTCACTATATGGAGGAGCGAGAG 2439  
Qy 4606 CTGTCTGTGATCGAGTAGCTATCATGTGCTTGGCAGTTAGATGTATCGGACAGTAC 4665  
Db 2440 CTGTCTGTGATCGAGTAGCTATCATGTGCTTGGCAGTTAGATGTATCGGACAGTAC 2499  
Qy 4666 AACATCTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTAATTTCAAGGACTGGA 4725  
Db 2500 AACATCTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTAATTTCAAGGACTGGA 2559  
Qy 4726 TAGAAAACTTAGAAGTAGACCCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAA 4785  
Db 2560 TAGAAAACTTAGAAGTAGACCCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAA 2619  
Qy 4786 GCCTTCAGGAAAGCTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGT 4845  
Db 2620 GCCTTCAGGAAAGCTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGT 2679  
Qy 4846 CCCTTTCAACATCTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTGCATTTGAAGAT 4905  
Db 2680 CCCTTTCAACATCTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTGCATTTGAAGAT 2739  
Qy 4906 ATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGG 4965  
Db 2740 ATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGG 2799  
Qy 4966 AGGAAGATAATAGTTGTGGAACCTTTTAAACAGCACACTTTTGGTGGGAACGAAACAGAGAG 5025  
Db 2800 AGGAAGATAATAGTTGTGGAACCTTTTAAACAGCACACTTTTGGTGGGAACGAAACAGAGAG 2859  
Qy 5026 ATAGAGTAGTATTTTGAATTTGTTGTTGCTGCTTACTGGGACTTTCTTTCTTTTC 5085  
Db 2860 ATAGAGTAGTATTTTGAATTTGTTGTTGCTGCTTACTGGGACTTTCTTTCTTTTC 2919  
Qy 5086 ACTTAATTTTAACTTTGGTTTAAAAAGTTTTTTTATTTGGAATGGTAACTCGAGAACCAAGA 5145  
Db 2920 ACTTAATTTTAACTTTGGTTTAAAAAGTTTTTTTATTTGGAATGGTAACTCGAGAACCAAGA 2979  
Qy 5146 ACCCAGCTTGAATTTTCTAAGCTCCTTAATTTGAATGCTGCTGTTGCTGTTTCTTTTC 5205  
Db 2980 ACCCAGCTTGAATTTTCTAAGCTCCTTAATTTGAATGCTGCTGTTGCTGTTTCTTTTC 3039  
Qy 5206 TCTTTAAATAAAACGATATGATAAATTAAGTGAAGCTGCACTGTTTGTATTTGAAGTATATTG 5265  
Db 3040 TCTTTAAATAAAACGATATGATAAATTAAGTGAAGCTGCACTGTTTGTATTGAAAGTATATTG 3099

5266 AACTATATAGTTGGTATGTCATCTTTTACCAATTCAGAAACAGTGTCTTGAATTTGTG 5325  
 3100 AACTATATAGTTGGTATGTCATCTTTTACCAATTCAGAAACAGTGTCTTGAATTTGTG 3159  
 5326 ATTAAAGGAATGTAATAGATAGTATTTATTTTAAAGTATCTTTAGTTTATGCCATC 5385  
 3160 ATTAAAGGAATGTAATAGATAGTATTTATTTTAAAGTATCTTTAGTTTATGCCATC 3219  
 5386 TTCTTAAATAGTACGTAATGTCCTCAATCTCAATTAATAAATACTAAT 5430  
 3220 TTCTTAAATAGTACGTAATGTCCTCAATCTCAATTAATAAATACTAAT 3264

RESULT 12  
 AX537472  
 LOCUS AX537472 4917 bp DNA linear PAT 23-NOV-2002  
 DEFINITION Sequence 3 from Patent WO02070690.  
 ACCESSION AX537472  
 VERSION AX537472.1 GI:25269279  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 Chen, H., Kilinski, L. and le Bihan, S.  
 Abca5 transporter and uses thereof  
 Patent: WO 02070690-A 3 12-SEP-2002;  
 Active Pass Pharmaceuticals, Inc. (CA)  
 Location/Qualifiers  
 1..4917  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 49.6%; Score 2713; DB 6; Length 4917;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2763; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 114 ATGTCCACTGCAATAGGAGGTAGGATTTGGAGACAGACAGACACTTCTACTGAAG 173  
 DB 1 ATGTCCACTGCAATAGGAGGTAGGATTTGGAGACAGACAGACACTTCTACTGAAG 60  
 QY 174 AATTACTTAATTAATGCAAGACCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 233  
 DB 61 AATTACTTAATTAATGCAAGACCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 120  
 QY 234 TTTTCTTTTATTTGGTTAATTAATTAATGCAAGACCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 293  
 DB 121 TTTTCTTTTATTTGGTTAATTAATTAATGCAAGACCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 180  
 QY 294 GTGCCCTAATAGAACTCAATCTATGCAAGTGTCTTCTTCTAATCTAATTTCTTGA 353  
 DB 181 GTGCCCTAATAGAACTCAATCTATGCAAGTGTCTTCTTCTAATCTAATTTCTTGA 240  
 QY 354 TATCTCCAGTGAATTAATTAATGCAAGACCAATGCAAGTGTCTTCTAATCTAATTTCTTGA 413  
 DB 241 TATCTCCAGTGAATTAATTAATGCAAGACCAATGCAAGTGTCTTCTAATCTAATTTCTTGA 300  
 QY 414 CTGTGATGCTAATTAATTAATGCAAGACCAATGCAAGTGTCTTCTAATCTAATTTCTTGA 473  
 DB 301 CTGTGATGCTAATTAATTAATGCAAGACCAATGCAAGTGTCTTCTAATCTAATTTCTTGA 360  
 QY 474 TCTAAGCCGAGCAACTTTGTAGTGTGTTTCAAGAATCTCAATGCTCTATGAATCTTGT 533  
 DB 361 TCTAAGCCGAGCAACTTTGTAGTGTGTTTCAAGAATCTCAATGCTCTATGAATCTTGT 420  
 QY 534 TTTTCTTCTGATGATGATCCAGTATCTTCTATTTATATGATTTCAAGAGCTGGCTGTCA 593  
 DB 421 TTTTCTTCTGATGATGATCCAGTATCTTCTATTTATATGATTTCAAGAGCTGGCTGTCA 480

QY 594 AATCATGTGAGGTGCTCAGTACTGGTCTCAGGTTTCCAGTCTTACAGTCTTACAGCATCCATA 653  
 DB 481 AATCATGTGAGGTGCTCAGTACTGGTCTCAGGTTTCCAGTCTTACAGTCTTACAGCATCCATA 540  
 QY 654 GATGCTGCCATTTATACAGTTGAAGCAATGTTTCTTTTGAAGGAGCTGGAGTCAACT 713  
 DB 541 GATGCTGCCATTTATACAGTTGAAGCAATGTTTCTTTTGAAGGAGCTGGAGTCAACT 600  
 QY 714 AAGCTGTTTATTATGAGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGGAGTA 773  
 DB 601 AAGCTGTTTATTATGAGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGGAGTA 660  
 QY 774 ATTTTAAATATACCTAGTTATAGCAATTTTCCCTTTTGGATAGTCTTTTGGCAATTCATATC 833  
 DB 661 ATTTTAAATATACCTAGTTATAGCAATTTTCCCTTTTGGATAGTCTTTTGGCAATTCATATC 720  
 QY 834 GTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATTAATGGGACTTTCATGATATC 893  
 DB 721 GTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATTAATGGGACTTTCATGATATC 780  
 QY 894 GCCTTTTGGCTTTTCCCTGGGTTCTTCTATATACAAGTTTAAATTTTCTTATGCTCCCTTCTT 953  
 DB 781 GCCTTTTGGCTTTTCCCTGGGTTCTTCTATATACAAGTTTAAATTTTCTTATGCTCCCTTCTT 840  
 QY 954 ATGGCAGTCAATTCGCGACAGCTTCTTTTGTATTTTCTCAAAGTAGCAGCAATTTGTATATT 1013  
 DB 841 ATGGCAGTCAATTCGCGACAGCTTCTTTTGTATTTTCTCAAAGTAGCAGCAATTTGTATATT 900  
 QY 1014 CTGCTTTTTCCTTTTATGATATCATCTGTATTTTGTCTTAAATGCTGACCTCTT 1073  
 DB 901 CTGCTTTTTCCTTTTATGATATCATCTGTATTTTGTCTTAAATGCTGACCTCTT 960  
 QY 1074 TTTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTACTGTGCTTTTGGATTT 1133  
 DB 961 TTTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTACTGTGCTTTTGGATTT 1020  
 QY 1134 ATTTGGCTTATGATTAATCTCATAGAAAGTTTCCAAATCTGTAGTGTGCTTTTCAGT 1193  
 DB 1021 ATTTGGCTTATGATTAATCTCATAGAAAGTTTCCAAATCTGTAGTGTGCTTTTCAGT 1080  
 QY 1194 CTTTCTGTCTGCTACTTTTGTGATTTGATGATGACAGGTGATGATTTAGAAATTTT 1253  
 DB 1081 CTTTCTGTCTGCTACTTTTGTGATTTGATGATGACAGGTGATGATTTAGAAATTTT 1140  
 QY 1254 AATGAAGTGTCTTCAATTTCAAATTTGACTGAGGCCCATATCTCTAATTTATTAATTT 1313  
 DB 1141 AATGAAGTGTCTTCAATTTCAAATTTGACTGAGGCCCATATCTCTAATTTATTAATTT 1200  
 QY 1314 ATCATGCTCACACTTAATAGTATTTCTATGCTCTGCTGCTGCTATCTTCAATCAAGTC 1373  
 DB 1201 ATCATGCTCACACTTAATAGTATTTCTATGCTCTGCTGCTGCTATCTTCAATCAAGTC 1260  
 QY 1374 ATTTCCAGGGGAATTTGGCTTACGGAGATCATCTTTTATATTTTCTGAAGCTTTCATATTGG 1433  
 DB 1261 ATTTCCAGGGGAATTTGGCTTACGGAGATCATCTTTTATATTTTCTGAAGCTTTCATATTGG 1320  
 QY 1434 TCAAGAGCAGAAAGAAATTTATGAGGAGTTATCAGAGGCAATGTTAAATGGAATTTATGCT 1493  
 DB 1321 TCAAGAGCAGAAAGAAATTTATGAGGAGTTATCAGAGGCAATGTTAAATGGAATTTATGCT 1380  
 QY 1494 TTTAGTGAATTTATTTGAGCCAGTTTCTTCAAGAAATTTGTAGGAAAGAAAGCCATAAGAAAT 1553  
 DB 1381 TTTAGTGAATTTATTTGAGCCAGTTTCTTCAAGAAATTTGTAGGAAAGAAAGCCATAAGAAAT 1440  
 QY 1554 AGTGGTATTCAGAGACATACAGAAAGAGGTTAAATGTTGGGCTTTTGAATTTG 1613  
 DB 1441 AGTGGTATTCAGAGACATACAGAAAGAGGTTAAATGTTGGGCTTTTGAATTTG 1500  
 QY 1614 TCAATTTGACATATATGAGGCTCAGATTACTGCTTACTTTGGCCACAGTGGAAACAGAAAG 1673  
 DB 1501 TCAATTTGACATATATGAGGCTCAGATTACTGCTTACTTTGGCCACAGTGGAAACAGAAAG 1560  
 QY 1674 AGTACATTTGATGAATTAATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1733











Db 481 TTGGATGGACAAGCTGTGTTGATATCCCTTATTTTTTATCATCTTATTTTGTGCT 540  
Qy 3368 AGGAAGCTTATGGCAATTCATATGGAATATATTTTATCTGTAAAGTTCCTTGCTGT 3427  
Db 541 AGGAAGCTTATGGCAATTCATATGGAATATATTTTATCTGTAAAGTTCCTTGCTGT 600  
Qy 3428 GGTGTTTTGCTTTATGTTTATGTTTCCATCAGTATTTCTGTTCACTATATGTTCTTT 3487  
Db 601 GGTGTTTTGCTTTATGTTTATGTTTCCATCAGTATTTCTGTTCACTATATGTTCTTT 660  
Qy 3488 CACCTTTAAGAAAATTTTAAATACCAAGAATTTGGTCATTTATCTATCTGTGGCAGC 3547  
Db 661 CACCTTTAAGAAAATTTTAAATACCAAGAATTTGGTCATTTATCTATCTGTGGCAGC 720  
Qy 3548 GTTGCTTGTATGCAATCAGTAAATAAATCTTTTATGGAATACCAATGCAACTAT 3607  
Db 721 GTTGCTTGTATGCAATCAGTAAATAAATCTTTTATGGAATACCAATGCAACTAT 780  
Qy 3608 TCTCAATATGCTTTGATCATCATTCCTCAATCTATCCATCTTAGGTTGCCTGATTC 3667  
Db 781 TCTCAATATGCTTTGATCATCATTCCTCAATCTATCCATCTTAGGTTGCCTGATTC 840  
Qy 3668 TTTTCATAAGATTTCTTGGAGAATGTACGAAAATGTGACACCTATATATCCATGGGA 3727  
Db 841 TTTTCATAAGATTTCTTGGAGAATGTACGAAAATGTGACACCTATATATCCATGGGA 900  
Qy 3728 TAGGCTTTCAAGTGTGTTATATCGCTTACCTGCAAGTGTGATGTTTCTCTTT 3787  
Db 901 TAGGCTTTCAAGTGTGTTATATCGCTTACCTGCAAGTGTGATGTTTCTCTTT 960  
Qy 3788 ACATATCTATGAGAAAATATGAGGAGATCAATTAAGAAAGATCCCTTTTTCAGAAA 3847  
Db 961 ACATATCTATGAGAAAATATGAGGAGATCAATTAAGAAAGATCCCTTTTTCAGAAA 1020  
Qy 3848 CCTTTCAAGAGTCTTAAATATGAGAGCTTCCAGAACCCAGACCAATGAGGATGAAGA 3907  
Db 1021 CCTTTCAAGAGTCTTAAATATGAGAGCTTCCAGAACCCAGACCAATGAGGATGAAGA 1080  
Qy 3908 TGAAGATGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3967  
Db 1081 TGAAGATGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1140  
Qy 3968 GGAGAAACCATCATTTATGCTCAGCAATTTGCATAAAGAAATATGATCAAGAAAGATTT 4027  
Db 1141 GGAGAAACCATCATTTATGCTCAGCAATTTGCATAAAGAAATATGATCAAGAAAGATTT 1200  
Qy 4028 TCTTCTTCAAGAAAGTAAAGAGTGGCAATCAATATCATCTCTTCTTCTGTGTGAAGAAA 4087  
Db 1201 TCTTCTTCAAGAAAGTAAAGAGTGGCAATCAATATCATCTCTTCTTCTGTGTGAAGAAA 1260  
Qy 4088 AGGAGAGATCTTAGGACTATGGGTCCAAATGGTCTGGCAAAAGCAAAATATTAATAT 4147  
Db 1261 AGGAGAGATCTTAGGACTATGGGTCCAAATGGTCTGGCAAAAGCAAAATATTAATAT 1320  
Qy 4148 TCTGTTTGTGATATTTGAACCAATCTCAGGCAAGTATTTTATGAGATATTTCTTCAAGA 4207  
Db 1321 TCTGTTTGTGATATTTGAACCAATCTCAGGCAAGTATTTTATGAGATATTTCTTCAAGA 1380  
Qy 4208 GACAAGTGAAGATGATGATTTCACTGAAGTGTATGGTGTACTGCTCCTCAGATAAACCCCTTT 4267  
Db 1381 GACAAGTGAAGATGATGATTTCACTGAAGTGTATGGTGTACTGCTCCTCAGATAAACCCCTTT 1440  
Qy 4268 GTGGCAGATCTATCAATTCAGGAACATTTTGAATTTTATGAGGCTGTCAAGGAATGAG 4327  
Db 1441 GTGGCAGATCTATCAATTCAGGAACATTTTGAATTTTATGAGGCTGTCAAGGAATGAG 1500  
Qy 4328 TGCAGTGCATGAAGAGTCAATAGTCAATACATGACATGCTGATTTTAAAGAAACA 4387  
Db 1501 TGCAGTGCATGAAGAGTCAATAGTCAATACATGACATGCTGATTTTAAAGAAACA 1560  
Qy 4388 TCTTCAGAGACTGTAAAGAAACATCTCCTGAGGAATCAACGAAAGTGTGTTTGTCTCT 4447

Db 1561 TCTTCAGAGACTGTAAAGAAACATCTCCTGAGGAATCAACGAAAGTGTGTTTGTCTCT 1620  
Qy 4448 AAGTATGCTAGGAAATCCTCAGATTTACTTTGTAGATGAACCATCTACAGGTATGATCC 4507  
Db 1621 AAGTATGCTAGGAAATCCTCAGATTTACTTTGTAGTGAACCATCTACAGGTATGATCC 1680  
Qy 4508 CAAAGCCAAACAGCAGCATGTGGCGAGCAATTCGAATCTGCAATTTAAACAGAAAGCGGC 4567  
Db 1681 CAAAGCCAAACAGCAGCATGTGGCGAGCAATTCGAATCTGCAATTTAAACAGAAAGCGGC 1740  
Qy 4568 TGTATTTCTGACCACTCACTATATGAGGAGGAGCAGAGGCTGTCTGTATCGAGTACGTAT 4627  
Db 1741 TGTATTTCTGACCACTCACTATATGAGGAGGAGCAGAGGCTGTCTGTATCGAGTACGTAT 1800  
Qy 4628 CATGTTGTCTGGCAGTTAAGATGTATCGGAACAGTACAACTAAAGAGTAAATTTGG 4687  
Db 1801 CATGTTGTCTGGCAGTTAAGATGTATCGGAACAGTACAACTAAAGAGTAAATTTGG 1860  
Qy 4688 AAAAGGCTACTTTTGGAAATTTAAATTTGAAGACCTGATAGAAAACCTAGAAAGTACCG 4747  
Db 1861 AAAAGGCTACTTTTGGAAATTTAAATTTGAAGACCTGATAGAAAACCTAGAAAGTACCG 1920  
Qy 4748 CTTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAAGAAAGTTTTCTTC 4807  
Db 1921 CTTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAAGAAAGTTTTCTTC 1980  
Qy 4808 TATTTTGGCTTATAAAATTTCTTAAGGAAGATGTTCACTCCCTTCACTCAATCTTTTTTAA 4867  
Db 1981 TATTTTGGCTTATAAAATTTCTTAAGGAAGATGTTCACTCCCTTCACTCAATCTTTTTTAA 2040  
Qy 4868 GCTGGAAGAGCTTAAACATGCTTTTCCCAATGGAAGAAATATAGCTTTTCAAGCAACAT 4927  
Db 2041 GCTGGAAGAGCTTAAACATGCTTTTCCCAATGGAAGAAATATAGCTTTTCAAGCAACAT 2100  
Qy 4928 GGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGGAGATTAATAGTTGTGAAC 4987  
Db 2101 GGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGGAGATTAATAGTTGTGAAC 2160  
Qy 4988 TTTTAAACAGCACACTTTTGTGGAAAGCAACAAAGAAAGATAGAGTAGTATTTGAATTTG 5047  
Db 2161 TTTTAAACAGCACACTTTTGTGGAAAGCAACAAAGAAAGATAGAGTAGTATTTGAATTTG 2220  
Qy 5048 TATGTTTGGTCTGCTTACTGGAATCTTCTTTTCACTTAAATTTTAACTTTGTTTAA 5107  
Db 2221 TATGTTTGGTCTGCTTACTGGAATCTTCTTTTCACTTAAATTTTAACTTTGTTTAA 2280  
Qy 5108 AAAAGTTTTTTTATTTGGAATGTTAACTGGAGAACCAAGAACGCACTTGAATTTTCTAAG 5167  
Db 2281 AAAAGTTTTTTTATTTGGAATGTTAACTGGAGAACCAAGAACGCACTTGAATTTTCTAAG 2340  
Qy 5168 CTCTTAAATGAAATGCTGTGTTGTTGTTTCTTTTCTTTTAAATAAAACGTATGTAT 5227  
Db 2341 CTCTTAAATGAAATGCTGTGTTGTTGTTTCTTTTCTTTTAAATAAAACGTATGTAT 2400  
Qy 5228 AATTAAGTGAAGTGCATGTTTGTATTTGAAGTATATTTGAATATATAGTTTGTATGTCAT 5287  
Db 2401 AATTAAGTGAAGTGCATGTTTGTATTTGAAGTATATTTGAATATATAGTTTGTATGTCAT 2460  
Qy 5288 CTTTTTCAACATTCAGAAACAGTCTGAAATTTGTGATTTTAAAGAAATTTGAATAGAA 5347  
Db 2461 CTTTTTCAACATTCAGAAACAGTCTGAAATTTGTGATTTTAAAGAAATTTGAATAGAA 2520  
Qy 5348 TAGTTTTTATTTTAAAGTATCTTTTAAAGTTTATGCCATCTTCTTTAAATAAGTACGTAATGT 5407  
Db 2521 TAGTTTTTATTTTAAAGTATCTTTTAAAGTTTATGCCATCTTCTTTAAATAAGTACGTAATGT 2580  
Qy 5408 TCCAATCTAAATAAAAACTAAATTCATACTAAATGATGAGAAAGATACATAAAGCAATG 5467  
Db 2581 TCCAATCTAAATAAAAACTAAATTCATACTAAATGATGAGAAAGATACATAAAGCAATG 2640  
Qy 5468 TGAAGTT 5475  
Db 2641 TGAAGTT 2648

RESULT 15  
 HSM807592 5243 bp mRNA linear PRI 30-AUG-2003  
 LOCUS Homo sapiens mRNA; cDNA DKFZp451f117 (from clone DKFZp451f117).  
 DEFINITION BX647447  
 ACCESSION BX647447  
 VERSION 1  
 KEYWORDS Homo sapiens (human)  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 5243)  
 Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,  
 Fobo, G., Han, M. and Wiemann, S.  
 The German Human cDNA Consortium  
 Direct Submission  
 Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764  
 Neuherberg, GERMANY  
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
 sequenced by BMFZ (Biomedical Research Center at the  
 Heinrich-Heine-University, Dueseldorf/Germany) within the cDNA  
 sequencing consortium of the German Genome Project. This clone  
 (DKFZp451f117) is available at the RZPD in Berlin. Please contact  
 the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further  
 information about the clone and the sequencing project is available  
 at http://mips.gsf.de/proj/cDNA/.  
 Location/Qualifiers  
 1..5243  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="DKFZp451f117"  
 /tissue\_type="human skeletal muscle"  
 /clone\_lib="451 (synonym: hlec1). Vector pSport1; host  
 DH10B; sites NotI + SalI"  
 /dev\_stages="adult"  
 5201..5206  
 polyA\_signal  
 polyA\_site 5225  
 ORIGIN  
 Query Match 41.9%; Score 2294; DB 9; Length 5243;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 2394; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 3035 TACTATGTTTATCTTTACCTATATTAGTGAATATCATTTAGTAACCTACTATCTTTATCA 3094  
 DB 1 TACTATGTTTATCTTTACCTATATTAGTGAATATCATTTAGTAACCTACTATCTTTATCA 60  
 QY 3095 TTTAAATGAGTGAACCATCCAGATCTGGAGTACCCCATCTTTCAAGAAATCTGA 3154  
 DB 61 TTTAAATGAGTGAACCATCCAGATCTGGAGTACCCCATCTTTCAAGAAATCTGA 120  
 QY 3155 TATAGTTTTTAAATTTAGCTGTATTTTCAAGCAGCTTTGCTGGAATCATTTGTTACTGC 3214  
 DB 121 TATAGTTTTTAAATTTAGCTGTATTTTCAAGCAGCTTTGCTGGAATCATTTGTTACTGC 180  
 QY 3215 AATGCCACTTACTTTGGCCATCGAATATGCGAGATCATTAAGATCAAGCTTATCTCA 3274  
 DB 181 AATGCCACTTACTTTGGCCATCGAATATGCGAGATCATTAAGATCAAGCTTATCTCA 240  
 QY 3275 ACTTAAACTTTTCAGCTCTTTTGGCCATCGAATATGCGAGATCATTAAGATCAAGCTTATCT 3334  
 DB 241 ACTTAAACTTTTCAGCTCTTTTGGCCATCGAATATGCGAGATCATTAAGATCAAGCTTATCT 300  
 QY 3335 CCCCTTATTTTATCATCTTATTTGATGCTAGGAGCTTATGGCAATTCATTTATGG 3394  
 DB 301 CCCCTTATTTTATCATCTTATTTGATGCTAGGAGCTTATGGCAATTCATTTATGG 360  
 QY 3395 ATTATATTTTATCATCTTAAAGTCTTGTGCTGTTTTCCTTATTTGCTTATTTGTTGCTCC 3454

DB 361 ATTATATTTTATCTGTAAAGTTCCTGTGCTGTTTTCCTTATTTGGTTATGTTCC 420  
 QY 3455 ATCAGTTATCTGTTCACTTATATTGCTTCTTACCCTTAAAGAAATTTTAAATACCA 3514  
 DB 421 ATCAGTTATCTGTTCACTTATATTGCTTCTTACCCTTAAAGAAATTTTAAATACCA 480  
 QY 3515 AGAATTTTGGTCAATTTATCTATCTGTTGGCAGCTTGGCTTATTTGCAATCACTGAAT 3574  
 DB 481 AGAATTTTGGTCAATTTATCTATCTGTTGGCAGCTTGGCTTATTTGCAATCACTGAAT 540  
 QY 3575 AACTTTCTTTATGGGATACACAATTTGCACTATTTCTTCAATATGCTTTTGTATCATCAT 3634  
 DB 541 AACTTTCTTTATGGGATACACAATTTGCACTATTTCTTCAATATGCTTTTGTATCATCAT 600  
 QY 3635 TCCAAATCTATCCACTTCTAGGTTGCTGATTTCTTTCATAAAGATTTCTTGAAGAATGT 3694  
 DB 601 TCCAAATCTATCCACTTCTAGGTTGCTGATTTCTTTCATAAAGATTTCTTGAAGAATGT 660  
 QY 3695 AGAAAAATTTGGACACCTATTAATCCATGGATAGGCTTTTCAGTAGCTTTTATATCGCC 3754  
 DB 661 AGAAAAATTTGGACACCTATTAATCCATGGATAGGCTTTTCAGTAGCTTTTATATCGCC 720  
 QY 3755 TTACCTGCACTGTGCTGATTTTCTTCAATATGAGAAAAATATGAGG 3814  
 DB 721 TTACCTGCACTGTGCTGATTTTCTTCAATATGAGAAAAATATGAGG 780  
 QY 3815 CAGATCAATPAGAAAGATCCCTTTTTCAGAAACCTTTCAACGAAAGTCTTAAATAAGAA 3874  
 DB 781 CAGATCAATPAGAAAGATCCCTTTTTCAGAAACCTTTCAACGAAAGTCTTAAATAAGAA 840  
 QY 3875 GCTTCAGAACCAACAGACCAATGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 3934  
 DB 841 GCTTCAGAACCAACAGACCAATGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 900  
 QY 3935 GGTCAAGAGCTGATGGTTGCGAGTTTGTGAGGAGAAACCATCCATTTATGTTGTCAGCAA 3994  
 DB 901 GGTCAAGAGCTGATGGTTGCGAGTTTGTGAGGAGAAACCATCCATTTATGTTGTCAGCAA 960  
 QY 3995 TTTGCAATAAGAAATGATGACAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGT 4054  
 DB 961 TTTGCAATAAGAAATGATGACAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGT 1020  
 QY 4055 GGCACAAATAATACATCTCTTTCTGTGTGAAAAAGAGAGATCTTATGAGACTATTTGGTCTCC 4114  
 DB 1021 GGCACAAATAATACATCTCTTTCTGTGTGAAAAAGAGAGATCTTATGAGACTATTTGGTCTCC 1080  
 QY 4115 AAATGGTGTGCAAAAGACCAATTTAATATTTCTGTTGGTGTATTTGAACCAACTTC 4174  
 DB 1081 AAATGGTGTGCAAAAGACCAATTTAATATTTCTGTTGGTGTATTTGAACCAACTTC 1140  
 QY 4175 AGCCAGGTATTTTATGAGATTTTCTTTCAGACCAAGTGAAGATGATGATTTCACTGAA 4234  
 DB 1141 AGCCAGGTATTTTATGAGATTTTCTTTCAGACCAAGTGAAGATGATGATTTCACTGAA 1200  
 QY 4235 GTGTATGGTTTACTGTCTCAGATAAACCCCTTTGTTGGCAGATATACATTTGCGAGAA 4294  
 DB 1201 GTGTATGGTTTACTGTCTCAGATAAACCCCTTTGTTGGCAGATATACATTTGCGAGAA 1260  
 QY 4295 TTTTGAATTTATGGAGCTGTCAAGGATGAGTGCAGATGACATGAAGAGTCAATAG 4354  
 DB 1261 TTTTGAATTTATGGAGCTGTCAAGGATGAGTGCAGATGACATGAAGAGTCAATAG 1320  
 QY 4355 TCGAATAACACATGCACTTTGATTTTAAAGAAACATCTTTCAGAAAGTGTAAAGAACTACC 4414  
 DB 1321 TCGAATAACACATGCACTTTGATTTTAAAGAAACATCTTTCAGAAAGTGTAAAGAACTACC 1380  
 QY 4415 TGCAGGAATCAACGAAAGTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4474  
 DB 1381 TGCAGGAATCAACGAAAGTGTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1440  
 QY 4475 TTTGCTAGATGAACCAATCTTACAGGTATGATCCCAAGCAACCAACCAACCAACCAACCAACCA 4534

DG 1441 TTTGCTAGATGAACCATCTACAGGTATGGATCCAAAGCCAAACAGCACATGTGGCGAGC 1500  
QY AATTCGAACTGCAATTTAAACAGAAAGCGGGCTGCTATTCTGACCACTCACTATATGGA 4594  
DB TTTTCTGAACTGCAATTTAAACAGAAAGCGGGCTGCTATTCTGACCACTCACTATATGGA 1560  
QY GGAGCGAGAGCTGCTGCTGATGCTGAGTAGCTATCATGCTGCTGGGAGTAAAGATGAT 4654  
DB GGAGCGAGAGCTGCTGCTGATGCTGAGTAGCTATCATGCTGCTGGGAGTAAAGATGAT 1620  
QY CGGAACACTACAAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAAT 4714  
DB CGGAACACTACAAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAAT 1680  
QY GAAGCACTGGATAGAAACCTAGAGTAGACCGGCTTCAAAGAGAAATTCAGTATATTT 4774  
DB GAAGCACTGGATAGAAACCTAGAGTAGACCGGCTTCAAAGAGAAATTCAGTATATTT 1740  
QY CCCAAATGCAAGCCGTCAGGAAAGTTTTCTTCTATTCTGCTTATTAATTCCTAAGGA 4834  
DB CCCAAATGCAAGCCGTCAGGAAAGTTTTCTTCTATTCTGCTTATTAATTCCTAAGGA 1800  
QY AGATGTTGAGTCCCTTTTCAAAATCTTTTTTAAAGCTGGAAGCTAAACATGCTTTGC 4894  
DB AGATGTTGAGTCCCTTTTCAAAATCTTTTTTAAAGCTGGAAGCTAAACATGCTTTGC 1860  
QY CATTCGAAGATATAGCTTTTCTCAAGCAACATTCGAACAGCTTTTCTAGAACTCACTAA 4954  
DB CATTCGAAGATATAGCTTTTCTCAAGCAACATTCGAACAGCTTTTCTAGAACTCACTAA 1920  
QY AGAACAAGAGGAGGAAGTAAATAGTTGTGGAACCTTTAAACAGCACACTTTGGTGGGAACG 5014  
DB AGAACAAGAGGAGGAAGTAAATAGTTGTGGAACCTTTAAACAGCACACTTTGGTGGGAACG 1980  
QY AACACAAGAGTATAGTATTTTGAATTTGATTTGCTGCTGCTTACTGGGACTT 5074  
DB AACACAAGAGTATAGTATTTTGAATTTGATTTGCTGCTGCTTACTGGGACTT 2040  
QY CTCTCTTTTCACTTAATTTAACTTTGTTTAAAGTTTATTTGGAATGGAATGTAACCTG 5134  
DB CTCTCTTTTCACTTAATTTAACTTTGTTTAAAGTTTATTTGGAATGGAATGTAACCTG 2100  
QY GAGAACCAAGAACGCACTGAAATTTTCTAAGCTCCTTAATGGAATGCTGCTGCTGCTG 5194  
DB GAGAACCAAGAACGCACTGAAATTTTCTAAGCTCCTTAATGGAATGCTGCTGCTGCTG 2160  
QY TGTCTTTGCTTTCTTTAAATTAAGCTATGATTAATTAAGTGAAGCTGCTGCTGCTGCT 5254  
DB TGTCTTTGCTTTCTTTAAATTAAGCTATGATTAATTAAGTGAAGCTGCTGCTGCTGCT 2220  
QY GAAGTATATGAACTATATAGTTTGTATGCTCATCTTTTCCACATTCAGAAACAGTCTT 5314  
DB GAAGTATATGAACTATATAGTTTGTATGCTCATCTTTTCCACATTCAGAAACAGTCTT 2280  
QY CTGAATTTGCTTTTAAAGGAAATGTAATAGATGTTTATTTTAAAGTATCTTTTAAAG 5374  
DB CTGAATTTGCTTTTAAAGGAAATGTAATAGATGTTTATTTTAAAGTATCTTTTAAAG 2340  
QY TTTATGCCATCTTCTTAAATTAAGTACGTAATGTTCCATCTTAAATTAAGTAAAT 5430  
DB TTTATGCCATCTTCTTAAATTAAGTACGTAATGTTCCATCTTAAATTAAGTAAAT 2396

Search completed: April 12, 2004, 15:37:03  
Job time : 13662 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2004, 07:51:45 ; Search time 1299 seconds

(without alignments)

17905.223 Million cell updates/sec

Title: US-10-090-458-4

Perfect score: 5475

Sequence: 1 gcgtccgcgcctccgcacag.....cataagcaatgtgaagtt 5475

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq 29Jan04.\*

1: geneseq1980s.\*

2: geneseq1990s.\*

3: geneseq2000s.\*

4: geneseq2001as.\*

5: geneseq2001bs.\*

6: geneseq2002s.\*

7: geneseq2003as.\*

8: geneseq2003bs.\*

9: geneseq2003cs.\*

10: geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5475	100.0	5475	7	AB57751 cDNA enco
2	5181	94.6	6525	6	ABN89594 Human ATP
3	5086	92.9	6369	6	AAD33648 Human TRI
4	4864	85.2	5262	6	AAD37620 Human tra
5	4572	83.5	4929	6	AAD37618 Human tra
6	4408	80.5	4785	6	AAD37619 Human tra
7	2826	51.6	5463	7	AB57749 cDNA enco
8	2739	50.0	3268	7	ADA53416 Human cod
9	2713	49.6	4917	7	AB57750 Coding se
10	2633	48.1	2723	9	ADC51606 Human mac
11	1497	27.3	1548	6	ABK3706 cDNA sequ
12	1350	24.7	2481	7	ABZ35938 Human sec
13	1343	24.5	3928	6	ABK3707 cDNA sequ
14	829	15.1	1632	7	ABZ35926 Human sec
15	783	14.3	1506	3	AAC81717 Human sec
16	744	13.6	1818	5	AA63176 Human pur
17	610	11.1	1677	5	AA570746 DNA enco
18	448	8.2	477	6	ABL87826 Human ova
19	440	8.0	612	6	ABL99789 Human sec
20	376	6.9	1346	6	ABL64459 Stomach c
21	376	6.9	1346	6	ABL63763 Breast ca
22	376	6.9	1346	6	ABN96909 Gene #340
23	358	6.5	389	3	AAA42471 Human sec

24	331	6.0	476	6	ABL81384	AB181384 Human ova
25	324	5.9	469	6	ABN95385	Gene #188
26	324	5.9	469	7	ABX74830	Human cDN
27	299	5.5	412	6	ABL81393	Human ova
28	286	5.2	1803	5	AA571497	DNA enco
29	284	5.2	431	6	ABL89460	Human pol
30	277	5.1	1994	6	ABN85542	Human CBP
31	243	4.4	2079	4	AA56391	ABC trans
32	242	4.4	341	2	AAV89159	EST clone
33	227	4.1	252	2	AAV89430	EST clone
34	213	3.9	7838	4	AA199209	Human exc
35	213	3.9	7838	4	AAK84993	Human imm
36	213	3.9	7838	5	AA163559	Human kid
37	200	3.7	440	5	ABV61636	Human pro
38	193	3.5	1478	4	AA199207	Human exc
39	193	3.5	1478	4	AAK84991	Human imm
40	193	3.5	1478	5	AA163557	Human kid
41	178	3.3	474	4	ABN8572	Human toe
42	178	3.3	474	4	AA138234	Probe #69
43	178	3.3	474	4	AAK32391	Human bon
44	178	3.3	474	4	AAK06692	Human bra
45	178	3.3	474	4	AB320398	Human liv

## ALIGNMENTS

### RESULT 1

AB57751 standard; cDNA; 5475 BP.

XX AB57751;

AC AB57751;

XX 04-FEB-2003 (first entry)

XX cDNA encoding novel human ATP binding cassette ABCA5 transporter #2.

XX Human; ATP binding cassette; ABC; ABCA5; transporter;

KW neurotxin transport; beta-amyloid peptide; chromosome mapping;

KW blood brain barrier transport; tissue typing; predictive medicine;

KW ABCA5 mediated disorder; ABCA5 related disorder; gene therapy; gene; ss.

XX Homo sapiens.

XX Key

XX Location/Qualifiers

FT 114..5042

FT /tag= a

FT /product= "ABCA5"

FT /note= "ATP binding cassette (ABC) A5 transporter"

XX US2002123107-A1.

XX 05-SEP-2002.

XX 01-MAR-2002; 2002US-00090458.

XX 02-MAR-2001; 2001US-0272885P.

XX (ACTI-) ACTIVE PASS PHARM INC.

XX Chen H, Kilinski L, Le Bihan S;

XX WPI; 2003-066798/06.

XX P-PSDB; ABG72424.

XX Novel isolated ATP binding cassette transporter family polypeptide,

PT ABCA5, useful for treating disorders associated with aberrant or unwanted

PT ABCA5 transporter expression or activity.

XX Claim 2; Page 39-42; 52pp; English.

XX The invention describes an isolated ATP binding cassette (ABC)

CC transporter family polypeptide (I), designated ABCA5. (I) or the

CC polynucleotide encoding it (II) are useful as targets for developing  
 CC modulating agents to regulate a variety of cellular processes,  
 CC particularly the transport of neurotoxic molecules, e.g., beta-amyloid  
 CC peptide (Aβ), across cell membranes or, e.g., the blood brain barrier  
 CC (BBB), as targets for developing modulating agents of multi-drug  
 CC resistance, as diagnostic and therapeutic tools, or to treat disorders  
 CC associated with aberrant or unwanted ABCA5 transporter expression or  
 CC activity. (I), (II) or a host cell (III) expressing (II) are useful in  
 CC screening assays, detection assays (e.g., chromosomal mapping, tissue  
 CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,  
 CC prognostic assays, monitoring clinical trials and pharmacogenomics), and  
 CC in methods of treatment (e.g., therapeutic and prophylactic). (I) or  
 CC (III) are useful as reagents or targets in assays applicable to treatment  
 CC and diagnosis of ABCA5-mediated or related disorders. (I) is useful to  
 CC screen for naturally occurring ABCA5 substrates; to screen for drugs or  
 CC compounds which modulate ABCA5 activity; as a bait protein in a yeast two  
 CC -hybrid or three-hybrid assay; and to identify other proteins which bind  
 CC to or interact with ABCA5. (II) is useful in: gene therapy; to detect  
 CC ABCA5 mRNA or a genetic alteration in a ABCA5 gene; to modulate ABCA5  
 CC activity; to locate gene regions associated with genetic disease or to  
 CC associate ABCA5 with the disease; to identify an individual from a minute  
 CC biological sample (tissue typing), and to aid in forensic identification  
 CC of the biological sample. This sequence encodes a novel human ATP binding  
 CC cassette (ABC) A5 transporter  
 XX  
 SQ Sequence 5475 BP; 1705 A; 917 C; 1048 G; 1804 T; 0 U; 1 Other;

Query Match 100.0%; Score 5475; DB 7; Length 5475;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 5475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 GGGTCGGCGCCCTCGCACAGATCCNAGCTGGTCCACCGCACTGAGTCAACAGACTCGAGC 60  
 1 GGGTCGGCGCCCTCGCACAGATCCNAGCTGGTCCACCGCACTGAGTCAACAGACTCGAGC 60  
 61 GGGTCAGGCGCTGACAGCTCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCT 120  
 61 GGGTCAGGCGCTGACAGCTCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCTCGGGCT 120  
 121 CTGCAATTAGGAGGTAGGAGTTGGAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 180  
 121 CTGCAATTAGGAGGTAGGAGTTGGAGACAGACAGACAGACAGACAGACAGACAGACAGACAG 180  
 181 TAATTAATGCAAGAACCAAGAGTAGTGTTCAGGAAATCTTTTCCACTATTTTTT 240  
 181 TAATTAATGCAAGAACCAAGAGTAGTGTTCAGGAAATCTTTTCCACTATTTTTT 240  
 241 TATTGTTGGTAAATTAATAGCATGATGCATCCAAATGAAGAAATGAAGAGTGCCTA 300  
 241 TATTGTTGGTAAATTAATAGCATGATGCATCCAAATGAAGAAATGAAGAGTGCCTA 300  
 301 ATATAGAACTCAATCCTATGGACAAAGTTTACTCTTTCTAACTAATCTTGGATATATCTC 360  
 301 ATATAGAACTCAATCCTATGGACAAAGTTTACTCTTTCTAACTAATCTTGGATATATCTC 360  
 361 CAGTGACTAATTAATCAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTCTGATG 420  
 361 CAGTGACTAATTAATCAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTCTGATG 420  
 421 TCATAATTAATCAAGAAATATACAAATGAAAGAAATGTTAACTCCAGTCTCTCTAAGC 480  
 421 TCATAATTAATCAAGAAATATACAAATGAAAGAAATGTTAACTCCAGTCTCTCTAAGC 480  
 481 CGAGCAACTTTGAGTGTGGTTTCAAGAGCTCCATGTCCTATGAACTCTGTTTTTTTC 540  
 481 CGAGCAACTTTGAGTGTGGTTTCAAGAGCTCCATGTCCTATGAACTCTGTTTTTTTC 540  
 541 CTGATATGATTCAGGATCTCTTATTTATGATTCAGGAGTGGCTGTTCAAAATCAT 600  
 541 CTGATATGATTCAGGATCTCTTATTTATGATTCAGGAGTGGCTGTTCAAAATCAT 600  
 601 GTGAGGCTGCTGAGTACTGGTCTCGAGTTTCAAGTTTCAAGTCTTCAAGTCCATAGATGCTG 660

Db 601 GTGAGGCTGCTGAGTACTGGTCTCAGGTTTCCAGTTTACAAAGCATCCATAGATGCTG 660  
 Qy 661 CCATTATACAGTTTGAAGACCAATGTTTCTTTTGGAGAGCTGGAGTCAACTTAAAGCTG 720  
 Db 661 CCATTATACAGTTTGAAGACCAATGTTTCTTTTGGAGAGCTGGAGTCAACTTAAAGCTG 720  
 Qy 721 TTATTATGGAGAAACTGCTGTGTGTAAGAAATAGATACCTTTTCCCGAGAGTAATTTAA 780  
 Db 721 TTATTATGGAGAAACTGCTGTGTGTAAGAAATAGATACCTTTTCCCGAGAGTAATTTAA 780  
 Qy 781 TATACCTAGTTATAGCATTTTCCACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG 840  
 Db 781 TATACCTAGTTATAGCATTTTCCACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG 840  
 Qy 841 AAAAGAAAGAAAGAAATATAAGAAATTTTAAAGATATATGGACTTCATGATACCTTTT 900  
 Db 841 AAAAGAAAGAAAGAAATATAAGAAATTTTAAAGATATATGGACTTCATGATACCTTTT 900  
 Qy 901 GGGTTTCTCGGGTCTCTCTATATACAGTTTAAATTTTCTTATGTCCTTCTTATGCGCAG 960  
 Db 901 GGGTTTCTCGGGTCTCTCTATATACAGTTTAAATTTTCTTATGTCCTTCTTATGCGCAG 960  
 Qy 961 TCATTGCGACAGCTTCTTTGTTTATTTCTTCAAGTAGCAGCATTTGATATTTCTGCTTT 1020  
 Db 961 TCATTGCGACAGCTTCTTTGTTTATTTCTTCAAGTAGCAGCATTTGATATTTCTGCTTT 1020  
 Qy 1021 TTTTCTCTTTATGCAATTCATCTGTATTTTTCCTTTTATGCTGACACCTCTTTTAA 1080  
 Db 1021 TTTTCTCTTTATGCAATTCATCTGTATTTTTCCTTTTATGCTGACACCTCTTTTAA 1080  
 Qy 1081 AATCAAAAACATGTGGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTTATTTGGCC 1140  
 Db 1081 AATCAAAAACATGTGGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTTATTTGGCC 1140  
 Qy 1141 TTATGATTAATCTCTATAGAAAGTTTTCCTCAATCTGTAGTGGCTTTTTCAGTCTTTCT 1200  
 Db 1141 TTATGATTAATCTCTCTATAGAAAGTTTTCCTCAATCTGTAGTGGCTTTTTCAGTCTTTCT 1200  
 Qy 1201 GTCACTGTACTTTTGTGATTTGTTGATTTGACAGCTCATGCAATTTAGAAGATTTAATGAAG 1260  
 Db 1201 GTCACTGTACTTTTGTGATTTGTTGATTTGACAGCTCATGCAATTTAGAAGATTTAATGAAG 1260  
 Qy 1261 GTGCTTCAATTTTCAAAATTTGACTGAGGCGCCATATCTCTAATTTATTAATATCATG 1320  
 Db 1261 GTGCTTCAATTTTCAAAATTTGACTGAGGCGCCATATCTCTAATTTATTAATATCATG 1320  
 Qy 1321 TCACACTTAATAGTATATTTCTCTTGTGCTGTCTATCTTGTATCAAGTCAATTCAG 1380  
 Db 1321 TCACACTTAATAGTATATTTCTCTTGTGCTGTCTATCTTGTATCAAGTCAATTCAG 1380  
 Qy 1381 GGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCATATTTGGTCAAAGA 1440  
 Db 1381 GGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCATATTTGGTCAAAGA 1440  
 Qy 1441 GCAAGAAATATAGAGGAGTTATCAGAGGCAATGTTAATGAAATATTTAGTTTATG 1500  
 Db 1441 GCAAGAAATATAGAGGAGTTATCAGAGGCAATGTTAATGAAATATTTAGTTTATG 1500  
 Qy 1501 AAATTTATGAGCCAGTTTCTTTCAGAAATTTGTAGGAAAGAACCAATTAAGATTTAGTGA 1560  
 Db 1501 AAATTTATGAGCCAGTTTCTTTCAGAAATTTGTAGGAAAGAACCAATTAAGATTTAGTGA 1560  
 Qy 1561 TTCAAGAGCATACAGAAAGAGGTTGAAGTGTGGAGCTTTGAGAAATTTGTCTATTG 1620  
 Db 1561 TTCAAGAGCATACAGAAAGAGGTTGAAGTGTGGAGCTTTGAGAAATTTGTCTATTG 1620  
 Qy 1621 ACATATATGAGGCTCAGATTTACTGCTTACTTTGGCCACAGTGGAAACAGAAAGAGTACAT 1680  
 Db 1621 ACATATATGAGGCTCAGATTTACTGCTTACTTTGGCCACAGTGGAAACAGAAAGAGTACAT 1680  
 Qy 1681 TGATGATATTTCTTGTGACTCTGCCACTCTTCTGATGGTTCGATCTATATATGAG 1740  
 Db 1681 TGATGATATTTCTTGTGACTCTGCCACTCTTCTGATGGTTCGATCTATATATGAG 1740



QY	1741	ACGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTGGCAATTTGTCCAC	1800
DB	1741	ACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTGGCAATTTGTCCAC	1800
QY	1801	AGTTAGATATACACTTTTGATGCTTTTGACAGTAGAAGAAAAATTTATCAATTTTGGCTTCAA	1860
DB	1801	AGTTAGATATACACTTTTGATGCTTTTGACAGTAGAAGAAAAATTTATCAATTTTGGCTTCAA	1860
QY	1861	TCAAAGGGATACCCAGCCAAACAATATAATACAAAGAGTCGAGAAAGTTTTACTAGATTTAG	1920
DB	1861	TCAAAGGGATACCCAGCCAAACAATATAATACAAAGAGTCGAGAAAGTTTTACTAGATTTAG	1920
QY	1921	ACATGCAGACTATCAAGAGATAACCAAGCTTAAAPAAATTAAGTGGTGCATAAAAAGAAAGC	1980
DB	1921	ACATGCAGACTATCAAGAGATAACCAAGCTTAAAPAAATTAAGTGGTGCATAAAAAGAAAGC	1980
QY	1981	TGTCATTAGGAATTTGCTGTTCTTGGGAACCCAAAGATACCTGCTCTAGATGAACCAACAG	2040
DB	1981	TGTCATTAGGAATTTGCTGTTCTTGGGAACCCAAAGATACCTGCTCTAGATGAACCAACAG	2040
QY	2041	CTGGAATGGACCCCTGTTCTCGACATATTTGATGGAAATCTTTTAAATAACAGAAAAAGCCA	2100
DB	2041	CTGGAATGGACCCCTGTTCTCGACATATTTGATGGAAATCTTTTAAATAACAGAAAAAGCCA	2100
QY	2101	ATCCGGGTGACAGTGTTCAGTACTCATTTTCATGGAATGAAGCTGACATTTCTTCAGATAGGA	2160
DB	2101	ATCCGGGTGACAGTGTTCAGTACTCATTTTCATGGAATGAAGCTGACATTTCTTCAGATAGGA	2160
QY	2161	AAGCTGTGATATCAAGAGAAATGCTGAAATGTGTGGTTCCTCAATGTTCTCTCAAAAGTA	2220
DB	2161	AAGCTGTGATATCAAGAGAAATGCTGAAATGTGTGGTTCCTCAATGTTCTCTCAAAAGTA	2220
QY	2221	AATCGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGTCGCACAGAAATCTC	2280
DB	2221	AATCGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGTCGCACAGAAATCTC	2280
QY	2281	TTTCTTCACCTGGTTTAAACAAACATATACCTGGAGCTACTTTTATCAACAGAAATGACCAAC	2340
DB	2281	TTTCTTCACCTGGTTTAAACAAACATATACCTGGAGCTACTTTTATCAACAGAAATGACCAAC	2340
QY	2341	AACCTGTGTATAGCTTGGCTTTCAGGACATGGACAAATTTTCAGGTTTTGTTTCTGCCCC	2400
DB	2341	AACCTGTGTATAGCTTGGCTTTCAGGACATGGACAAATTTTCAGGTTTTGTTTCTGCCCC	2400
QY	2401	TAGACAGTCATTCAAATTTGGGTGTCATTCTCTATGTGTGTTCCATCAGCACTTTTGGAG	2460
DB	2401	TAGACAGTCATTCAAATTTGGGTGTCATTCTCTATGTGTGTTCCATCAGCACTTTTGGAG	2460
QY	2461	ACGTAATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTTATAGTGTATTTA	2520
DB	2461	ACGTAATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTTATAGTGTATTTA	2520
QY	2521	CTCAGCAGCCACTGGAGGAGAAATGGAATCAAAATCTTTTGTAGTAATGGAAACAGAGCT	2580
DB	2521	CTCAGCAGCCACTGGAGGAGAAATGGAATCAAAATCTTTTGTAGTAATGGAAACAGAGCT	2580
QY	2581	TACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGGAAACAAC	2640
DB	2581	TACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGGAAACAAC	2640
QY	2641	AGATGTATACAAATGAGAAAGTTTCATTTCTTACCTTGAAACGTGAAGTAATCAGTGA	2700
DB	2641	AGATGTATACAAATGAGAAAGTTTCATTTCTTACCTTGAAACGTGAAGTAATCAGTGA	2700
QY	2701	GATCAGTGTGCTTCTGCTTTTAAATTTTTTTTTCAGATTCAGATTTTTATGTTTTGGTTC	2760
DB	2701	GATCAGTGTGCTTCTGCTTTTAAATTTTTTTTTCAGATTCAGATTTTTATGTTTTGGTTC	2760
QY	2761	ATCACTCTTTTAAAAATGCTGTGTTTCCCATCAAACCTTGTTCAGACCTTATATTTCTAA	2820
DB	2761	ATCACTCTTTTAAAAATGCTGTGTTTCCCATCAAACCTTGTTCAGACCTTATATTTCTAA	2820

QY	2821	AACCTGGAGACAACACATATAAATAACAAACAAGTCTGCTTCTTCAAAATTCCTGCTGACT	2880
DB	2821	AACCTGGAGACAACACATATAAATAACAAACAAGTCTGCTTCTTCAAAATTCCTGCTGACT	2880
QY	2881	CAGATATCAGTGTACTTATTAGCTTTTTCACAGCCAGAACATAAATGGTGACCATGATTA	2940
DB	2881	CAGATATCAGTGTACTTATTAGCTTTTTCACAGCCAGAACATAAATGGTGACCATGATTA	2940
QY	2941	ATGACAGTACTATGATACCGTGGCTCCCATAGTGGGCTTTAAATGTGATGCATTTCAG	3000
DB	2941	ATGACAGTACTATGATACCGTGGCTCCCATAGTGGGCTTTAAATGTGATGCATTTCAG	3000
QY	3001	AAAAGGACTATGTTTTTGGCAGCTGTTTTCAACAGTACTATGTTTTTAAATTCCTTTACCTATAT	3060
DB	3001	AAAAGGACTATGTTTTTGGCAGCTGTTTTCAACAGTACTATGTTTTTAAATTCCTTTACCTATAT	3060
QY	3061	TAGTGAATATCATTAGTAACTACTATCTTTATCATTTAAATGTGATGAAACCATCCAGA	3120
DB	3061	TAGTGAATATCATTAGTAACTACTATCTTTATCATTTAAATGTGATGAAACCATCCAGA	3120
QY	3121	TCCTGGATGCCCATCTTTTCAAGAAATTACTATAGTTTTTAAATTTGAGCTGTATT	3180
DB	3121	TCCTGGATGCCCATCTTTTCAAGAAATTACTATAGTTTTTAAATTTGAGCTGTATT	3180
QY	3181	TTCAAGCAGCTTTCCTTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCCATGGAAA	3240
DB	3181	TTCAAGCAGCTTTCCTTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCCATGGAAA	3240
QY	3241	ATGCAGAGAAATCATAGATCAAAAGCTTATACTCAACTTAAACCTTCAGGCTTTTGCCAT	3300
DB	3241	ATGCAGAGAAATCATAGATCAAAAGCTTATACTCAACTTAAACCTTCAGGCTTTTGCCAT	3300
QY	3301	CTGCATATTGGATTGGACAAGCTGTGTGATATCCCTTATTTTTTATCATTTCTTATTT	3360
DB	3301	CTGCATATTGGATTGGACAAGCTGTGTGATATCCCTTATTTTTTATCATTTCTTATTT	3360
QY	3361	TGATGCTPAGGAAGCTTATTGGCATTTCAATATGGAATATATTTTTTATCTGTHAAGTTCC	3420
DB	3361	TGATGCTPAGGAAGCTTATTGGCATTTCAATATGGAATATATTTTTTATCTGTHAAGTTCC	3420
QY	3421	TTGCTGTGGTTTTTTCGCTTATTTGGTTATGTTCCATCAGTTATCTTGTTGCACATTATTTG	3480
DB	3421	TTGCTGTGGTTTTTTCGCTTATTTGGTTATGTTCCATCAGTTATCTTGTTGCACATTATTTG	3480
QY	3481	CTTCTTTTCACTTTTAAAGAAATTTAAATACCAAAGAAATTTTGGTCATTATCTATCTTG	3540
DB	3481	CTTCTTTTCACTTTTAAAGAAATTTAAATACCAAAGAAATTTTGGTCATTATCTATCTTG	3540
QY	3541	TGGCAGCGTTGGCTGTGATTGCAATCACTGAAATAACTTCTTTATGGGATACACAATTG	3600
DB	3541	TGGCAGCGTTGGCTGTGATTGCAATCACTGAAATAACTTCTTTATGGGATACACAATTG	3600
QY	3601	CAACTATTCCTCATTTAGCCCTTTGTATCATCATTCCAATCTATCCACTTCCTAGGTTGCC	3660
DB	3601	CAACTATTCCTCATTTAGCCCTTTGTATCATCATTCCAATCTATCCACTTCCTAGGTTGCC	3660
QY	3661	TGATTTCTTTTCAATAAGATTTCTTGGAGAAATGTACGAAAAATGTGGACACCTATAATC	3720
DB	3661	TGATTTCTTTTCAATAAGATTTCTTGGAGAAATGTACGAAAAATGTGGACACCTATAATC	3720
QY	3721	CATGGGATAGCTTTTCAGTAGCTGTTATATCGCCTTACCTGCAAGTGTACTCTGGATTT	3780
DB	3721	CATGGGATAGCTTTTCAGTAGCTGTTATATCGCCTTACCTGCAAGTGTACTCTGGATTT	3780
QY	3781	TCCTCTTTCAATACTATGAGAAAAAATATGGAGGCAGATCAATPAGAAAAAGATCCCTTTT	3840
DB	3781	TCCTCTTTCAATACTATGAGAAAAAATATGGAGGCAGATCAATPAGAAAAAGATCCCTTTT	3840
QY	3841	TCAGAAACCTTTCAAACGAGTCTTAAAAATAGGAGCTTCCAGAACCAACAGCAATGAGG	3900
DB	3841	TCAGAAACCTTTCAAACGAGTCTTAAAAATAGGAGCTTCCAGAACCAACAGCAATGAGG	3900
QY	3901	ATGAAGATGAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGGTTGCCAGT	3960



Db 3901 ATGAAGATGAAGATGATCAAGCTGAAGAGCTAAGGCTCAAGAGCTGATGGTGGCCAGT 3960  
QY 3961 GTTGTGAGAGAGAAACCATCATTTATGGTCAGCAATTTGCAATAAGAAATATGATGACAGA 4020  
Db 3961 GTTGTGAGAGAGAAACCATCATTTATGGTCAGCAATTTGCAATAAGAAATATGATGACAGA 4020  
QY 4021 AAGATTTTCTTCTTCAAGAAAGTAAGAAAGTGGCAACTAAATACATCTCTTCTGTG 4080  
Db 4021 AAGATTTTCTTCTTCAAGAAAGTAAGAAAGTGGCAACTAAATACATCTCTTCTGTG 4080  
QY 4081 TGAAGAAAGAGAGATCTTAGGACTATTGGGTCCTAAGTGGTGGCAAGGCAAGTAA 4140  
Db 4081 TGAAGAAAGAGAGATCTTAGGACTATTGGGTCCTAAGTGGTGGCAAGGCAAGTAA 4140  
QY 4141 TTAATATTCTGTTGGTGATATTGAACCAACTTCAGGCCAGGATTTTATGAGAGATTAT 4200  
Db 4141 TTAATATTCTGTTGGTGATATTGAACCAACTTCAGGCCAGGATTTTATGAGAGATTAT 4200  
QY 4201 CTTTCAGAGCAAGTGAAGATGATGATTCAGTGAAGTGTGGTGTACGTCTCAGATAA 4260  
Db 4201 CTTTCAGAGCAAGTGAAGATGATGATTCAGTGAAGTGTGGTGTACGTCTCAGATAA 4260  
QY 4261 ACCCTTTGTGGCCAGATCACTATTGCAGGAACATTTTGAATTTATGGAGCTGTCAAAG 4320  
Db 4261 ACCCTTTGTGGCCAGATCACTATTGCAGGAACATTTTGAATTTATGGAGCTGTCAAAG 4320  
QY 4321 GAATGAGTCAAGTGCATGAAGAGTCAATAGTTCGAATACACATGCACTTGAATTAA 4380  
Db 4321 GAATGAGTCAAGTGCATGAAGAGTCAATAGTTCGAATACACATGCACTTGAATTAA 4380  
QY 4381 AAGAACATCTTCAGAGAGCTGTAAGAGAACTACCTGCAGGAATCAAAAGGAAAGTGTGTT 4440  
Db 4381 AAGAACATCTTCAGAGAGCTGTAAGAGAACTACCTGCAGGAATCAAAAGGAAAGTGTGTT 4440  
QY 4441 TTGCTCTAAGTATGTAGGAACTCTCAGATTACTTTGCTAGATGAACCATCTACAGGTA 4500  
Db 4441 TTGCTCTAAGTATGTAGGAACTCTCAGATTACTTTGCTAGATGAACCATCTACAGGTA 4500  
QY 4501 TGGATCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAACCTGATTTAAACAGAA 4560  
Db 4501 TGGATCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAACCTGATTTAAACAGAA 4560  
QY 4561 AGCGGCTGCTATCTTGACCACTACTATATGAGAGAGGCGAGGCTGTCTGTGATCGAG 4620  
Db 4561 AGCGGCTGCTATCTTGACCACTACTATATGAGAGAGGCGAGGCTGTCTGTGATCGAG 4620  
QY 4621 TAGCTATCATGTGTCTGGCAGTTAAGATGTATCGGAACAGTACACATCTAAGAGTA 4680  
Db 4621 TAGCTATCATGTGTCTGGCAGTTAAGATGTATCGGAACAGTACACATCTAAGAGTA 4680  
QY 4681 AATTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAACCTAGAAG 4740  
Db 4681 AATTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAACCTAGAAG 4740  
QY 4741 TAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATTCGAAGCGCTCAGAAAGTT 4800  
Db 4741 TAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATTCGAAGCGCTCAGAAAGTT 4800  
QY 4801 TTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAGATGTTTCAGTCCCTTTTCACAATCTT 4860  
Db 4801 TTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAGATGTTTCAGTCCCTTTTCACAATCTT 4860  
QY 4861 TTTTAAAGTGAAGAGCAATCAATGCTTTTGGCAATTAAGATATAGCTTTTCTCAAG 4920  
Db 4861 TTTTAAAGTGAAGAGCAATCAATGCTTTTGGCAATTAAGATATAGCTTTTCTCAAG 4920  
QY 4921 CAACATTGGAACAGGTTTTTGTAGAACTCACTAAGAGAAAGAGGAGGAGATATAGTT 4980  
Db 4921 CAACATTGGAACAGGTTTTTGTAGAACTCACTAAGAGAAAGAGGAGGAGATATAGTT 4980  
QY 4981 GTGGAACCTTTAAACAGCACATTTGGTGGGAAACGAAACAGAGATAGATGATTTT 5040

Db 4981 GTGGAACTTTAAACAGCACACTTTGGTGGGAAACGAAACACAAGAGATAGATGATTTT 5040  
QY 5041 GAATTTGTATTGTTGGTCTGCTTACTGGGACTTCTTTTCTTTTCACTTAATTTAACTT 5100  
Db 5041 GAATTTGTATTGTTGGTCTGCTTACTGGGACTTCTTTTCTTTTCACTTAATTTAACTT 5100  
QY 5101 TGGTTTAAAGAGTTTTTTTATTTGGAATGGTAACTGGAGAACCAAGAACGACCTTGAATTT 5160  
Db 5101 TGGTTTAAAGAGTTTTTTTATTTGGAATGGTAACTGGAGAACCAAGAACGACCTTGAATTT 5160  
QY 5161 TTCTAAGCTCCTTAAATGAAATGCTGTGGTGTGTTTTGTTTTTCTTTTAAATAAACG 5220  
Db 5161 TTCTAAGCTCCTTAAATGAAATGCTGTGGTGTGTTTTGTTTTTCTTTTAAATAAACG 5220  
QY 5221 TATGTATAATTAAGTGAAGCTGCATGTTTGTATTTGAAGTATATTGAATATATAGTTTGT 5280  
Db 5221 TATGTATAATTAAGTGAAGCTGCATGTTTGTATTTGAAGTATATTGAATATATAGTTTGT 5280  
QY 5281 ATGTCATCTTTTCCACCATTCAGAAACAGTCTCTGAATTTGTGATTTTAAAGGAATTTGT 5340  
Db 5281 ATGTCATCTTTTCCACCATTCAGAAACAGTCTCTGAATTTGTGATTTTAAAGGAATTTGT 5340  
QY 5341 AATAGAAATAGTTTATTTTAAAGTTATCTTTAAGTTTATGCCATCTTCTTAAATAGTAC 5400  
Db 5341 AATAGAAATAGTTTATTTTAAAGTTATCTTTAAGTTTATGCCATCTTCTTAAATAGTAC 5400  
QY 5401 GTAATGTTCCAACTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5460  
Db 5401 GTAATGTTCCAACTTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5460  
QY 5461 AGCAATGTGAAAGTT 5475  
Db 5461 AGCAATGTGAAAGTT 5475

RESULT 2  
ABN89594  
ID ABN89594 standard; cDNA; 6525 BP.  
XX ABN89594;  
XX AC AC  
XX ABN89594;  
DT 18-SEP-2002 (first entry)  
XX Human ATP-binding cassette transporter ABCA5 cDNA SEQ ID NO:1.  
DE Human; ABCA5; ABCA6; ABCA9; ABCA10; ATP-binding cassette transporter;  
XX chromosome 17; chromosome 17q; chromosome 17q24; antiarteriosclerotic;  
KW choletherapy; cholesterol; lipophilic molecule; inflammation;  
KW prostaglandin; prostacyclin; arteriosclerosis; transport; gene; ss.  
XX Homo sapiens.  
OS  
XX WO200246458-A2.  
PN  
XX 13-JUN-2002.  
PD  
XX 07-DEC-2001; 2001WO-RP015401.  
XX  
PF 07-DEC-2000; 2000EP-00403440.  
XX  
PR 23-JAN-2001; 2001US-0263231P.  
XX  
XX (AVET ) AVENTIS PHARMA SA.  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Denefle P, Rosier-Montus M, Prades C, Arnould-Reguigne I;  
PI Duverger N, Allikmets R, Dean M;  
XX P-PSDB; ABB81574.  
DR WPI; 2002-557584/59.  
XX  
XX A novel nucleic acid corresponding to ATP-binding cassette transporter  
PT genes and the encoded polypeptide, useful for preventing or treating a  
PT dysfunction in reverse transport of cholesterol.

Claim 1; Page 151-153; 216pp; English.

The present invention describes human ATP-binding cassette transporters (ABC). Specifically described are the human ABCA5, ABCA6, ABCA9 and ABCA10 genes (see ABN89594 to ABN9597) which encode the proteins given in ABN81574 to ABN81577). ABN89598 to ABN89715 represent ABCA5, ABCA6, ABCA9 and ABCA10 nucleotide fragments, and ABN89716 to ABN89806 represent primers for ABCA5, ABCA6, ABCA9 and ABCA10 genes which are used in the exemplification of the present invention. The ABC sequences have antiarteriosclerotic activities and can be used in gene therapy. ABC sequences can be used in the manufacture of a medicament intended for the prevention and/or treatment of a subject affected by a dysfunction in the reverse transport of cholesterol. The ABC proteins are involved in the reverse transport of cholesterol, in membrane transport of lipophilic molecules, in particular inflammation mediating substance such as prostaglandins and prostacyclins, or in any pathology whose candidate chromosomal region is situated on chromosome 17. They are also useful for the manufacture of a medicament intended for prevention of arteriosclerosis in various forms. The ABCA5, ABCA6, ABCA9 and ABCA10 genes are located to chromosome 17, more specifically to the 17q24 locus

XX PS CC  
XX CC  
XX CC

Query Match 94.6%; Score 5181; DB 6; Length 6525;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5331; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 97 AGGTTTATTCAGAAACATGTCCACTGCATTAGGGAGTAGGAGTTGGAGACAGACC A 156  
DB 994 AGGTTTATTCAGAAAACATGTCCACTGCCAATTAAGGGAGTAGGAGTTGGAGACAGACC A 1053  
QY 157 GAACACTCTCTACGAAGAATTACTTAATTAATGACAGAACCAAAAGAGTAGTGTCAGG 216  
DB 1054 GRACACTCTCTACGAAGAATTACTTAATTAATGACAGAACCAAAAGAGTAGTGTCAGG 1113  
QY 217 AAATCTTTTTTCCACTATTTTTTTTTTGGTTTAAATTAATTAATGACATGATGCATCAA 276  
DB 1114 AAATCTTTTTTCCACTATTTTTTTTTTGGTTTAAATTAATTAATGACATGATGCATCAA 1173  
QY 277 ATAAGAAATATGAAGAGTGCCCTAATATAGAACTCAATCCTATGACCAAGTTTACTCTTT 336  
DB 1174 ATAAGAAATATGAAGAGTGCCCTAATATAGAACTCAATCCTATGACCAAGTTTACTCTTT 1233  
QY 337 CTAATCTAATTTCTGGATATACTCCAGTGACTAATTAATTAACAAGCAGCATCATGAGAAAG 396  
DB 1234 CTAATCTAATTTCTGGATATACTCCAGTGACTAATTAATTAACAAGCAGCATCATGAGAAAG 1293  
QY 397 TGTCTACTGATCATCTACCTGATGTCATAATTAATTAACAAGATATACAAATGAAAAAGAA 456  
DB 1294 TGTCTACTGATCATCTACCTGATGTCATAATTAATTAACAAGATATATCAAAATGAAAAAGAA 1553  
QY 457 TGTTAAACATCCAGTCTCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAGACTCCA 516  
DB 1354 TGTTAAACATCCAGTCTCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAGACTCCA 1413  
QY 517 TGTCTATGAACCTTCGTTTTTTTCCGTATGATGATCCAGTATCTCTATTTATATGGAAT 576  
DB 1414 TGTCTATGAACCTTCGTTTTTTTCCGTATGATGATCCAGTATCTCTATTTATATGGAAT 1473  
QY 577 CAAGAGCTGGCTGTTCAAAATCATGTAGCGCTGCTCAGTACTGTGTCCTCAGGTTTTCAG 636  
DB 1474 CAAGAGCTGGCTGTTCAAAATCATGTAGCGCTGCTCAGTACTGTGTCCTCAGGTTTTCAG 1533  
QY 637 TTTTACAAGCATCCATAGATGCTGCATPATACATGATGAAGCACCAATGTTCTCTTTGGA 696  
DB 1534 TTTTACAAGCATCCATAGATGCTGCATPATATACATGATGAAGCACCAATGTTCTCTTTGGA 1593  
QY 697 AGGAGCTGGAGTCAACTAAGCTGTTTATATGCGGAGAACTGCTGTTGTAGAAATAGATA 756  
DB 1594 AGGAGCTGGAGTCAACTAAGCTGTTTATATGCGGAGAACTGCTGTTGTAGAAATAGATA 1653  
QY 757 CCTTTCCCCGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTCACTTTTGGATACT 816

2734 AAAATTTATCAATTTTGGCTTCAATCAAAAGGATACAGCCAAACATATATAACAAGAG 2793 Db  
1897 TCGAAGAGGTTTTACTAGATTTAGACATGACAGACTATCAAGATACCAAGCTAAAAAT 1956 Qy  
2794 TCGAAGAGGTTTTACTAGATTTAGACATGACAGACTATCAAGATACCAAGCTAAAAAT 2853 Db  
1957 TAAGTGGTGGTCAAAAAGAGCTGCTCATTTAGGAATGCTGTTCTGGGAACCAAGA 2016 Qy  
2854 TAAGTGGTGGTCAAAAAGAGCTGCTCATTTAGGAATGCTGTTCTGGGAACCAAGA 2913 Db  
2017 TACTGCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATGATGGA 2076 Qy  
2914 TACTGCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATGATGGA 2973 Db  
2077 ATCTTTTAAATACAGAAAGCCATCGGTGACAGTGTTCAGTACTCATTTTCATGATG 2136 Qy  
2974 ATCTTTTAAATACAGAAAGCCATCGGTGACAGTGTTCAGTACTCATTTTCATGATG 3033 Db  
2137 AAGCTGACATCTTCGACATAGGAAGCTGTGATATCAAGAGGAATGCTGAAATGCTGTG 2196 Qy  
3034 AAGCTGACATCTTCGACATAGGAAGCTGTGATATCAAGAGGAATGCTGAAATGCTGTG 3093 Db  
2197 GTTCTTCAATGTTCTCCTAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256 Qy  
3094 GTTCTTCAATGTTCTCCTAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153 Db  
2257 ACAAAATATTGCCCACAGAACTCTTCTTCACTGTTTAAACAAACATATACCTGGAGCTA 2316 Qy  
3154 ACAAAATATTGCCCACAGAACTCTTCTTCACTGTTTAAACAAACATATACCTGGAGCTA 3213 Db  
2317 CTTTATACAGAGAAATGACCAACCTGTGTGATATAGCTTGCCTTTCAAGACATGGACA 2376 Qy  
3214 CTTTATACAGAGAAATGACCAACCTGTGTGATATAGCTTGCCTTTCAAGACATGGACA 3273 Db  
2377 AATTTTCAGTTTGTGTTCTGCGCTAGACAGTCAATCAAAATGGGTGTCATTTCTTATG 2436 Qy  
3274 AATTTTCAGTTTGTGTTCTGCGCTAGACAGTCAATCAAAATGGGTGTCATTTCTTATG 3333 Db  
2437 GTGTTTCCATGACGCTTTTGGAGACGTAATTTTAAAGCTAGAGTTGAACAGAAATTTG 2496 Qy  
3334 GTGTTTCCATGACGCTTTTGGAGACGTAATTTTAAAGCTAGAGTTGAACAGAAATTTG 3393 Db  
2497 ACCAAGCAGATTAATAGTGTATTTACTCAGACGCCACTGGAGGAAGAAATGATCAAAAT 2556 Qy  
3394 ACCAAGCAGATTAATAGTGTATTTACTCAGACGCCACTGGAGGAAGAAATGATCAAAAT 3453 Db  
2557 CTTTGTGAAATGGAACAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGA 2616 Qy  
3454 CTTTGTGAAATGGAACAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGA 3513 Db  
2617 GCACCATGAGCCTTTTGGAAACAGAGTATACAAATAGCAAGTTTCATTTCTTTACCT 2676 Qy  
3514 GCACCATGAGCCTTTTGGAAACAGAGTATACAAATAGCAAGTTTCATTTCTTTACCT 3573 Db  
2677 TGAACCGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736 Qy  
3574 TGAACCGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 3633 Db  
2737 TTCAGATTTTATGTTTTTGGTTTCACTCTTTTAAATGCTGTGTTCCCATCAAC 2796 Qy  
3634 TTCAGATTTTATGTTTTTGGTTTCACTCTTTTAAATGCTGTGTTCCCATCAAC 3693 Db  
2797 TTTGTTCCAGACTTATTTTCTTAAACCTGGAGACAAACCAATAAATACAAACAAAGTC 2856 Qy  
3694 TTTGTTCCAGACTTATTTTCTTAAACCTGGAGACAAACCAATAAATACAAACAAAGTC 3753 Db  
2857 TGCCTTCTCAAATTCCTGCTGACTCAGATATCAGTGTCTTATTTAGCTTTTTCACAGCC 2916 Qy  
3754 TGCCTTCTCAAATTCCTGCTGACTCAGATATCAGTGTCTTATTTAGCTTTTTCACAGCC 3813 Db  
2917 AGAACATATGCTGACGATGATTAATGACAGTACTATGATTCGTGCTGCCCATAGTG 2976 Qy  
3814 AGAACATATGCTGACGATGATTAATGACAGTACTATGATTCGTGCTGCCCATAGTG 3873 Db

2977 CGGCTTTAAATGATGATGCAATTCAGAAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTA 3036 Qy  
3874 CGGCTTTAAATGATGATGCAATTCAGAAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTA 3933 Db  
3037 CTATGCTTTTATTTCTTTACCTATATTAGTGAATATCATTTAGTAACACTACTATCTTTTCAAT 3096 Qy  
3934 CTATGCTTTTATTTCTTTACCTATATTAGTGAATATCATTTAGTAACACTACTATCTTTTCAAT 3993 Db  
3097 TAAATGTGACTGAACCAATCCAGATCTGGAGTACCCCAATCTTTTCAAGAAATATCTGATA 3156 Qy  
3994 TAAATGTGACTGAACCAATCCAGATCTGGAGTACCCCAATCTTTTCAAGAAATATCTGATA 4053 Db  
3157 TAGTTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCA 3216 Qy  
4054 TAGTTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCA 4113 Db  
3217 TGGCACTTTACTTTGCTGATGGAATGCAAGAAATCATAAGATCAAAAGCTTATATCTCAAC 3276 Qy  
4114 TGGCACTTTACTTTGCTGATGGAATGCAAGAAATCATAAGATCAAAAGCTTATATCTCAAC 4173 Db  
3277 TTAACCTTTTCAAGCTCTTTTCCCATCTGCATATTGGAATGGACAGCTGTTGTTGATATCC 3336 Qy  
4174 TTAACCTTTTCAAGCTCTTTTCCCATCTGCATATTGGAATGGACAGCTGTTGTTGATATCC 4233 Db  
3337 CCTTATTTTATCATCTTTTATTTTATGCTAGGAAGCTTATTTGCAATTTTCATTTATGAT 3396 Qy  
4234 CCTTATTTTATCATCTTTTATTTTATGCTAGGAAGCTTATTTGCAATTTTCATTTATGAT 4293 Db  
3397 TATATTTTATGCTGTAAGTTTCTTCTGCTGCTGCTTTTGGCTTATTTGGTATGTTCCAT 3456 Qy  
4294 TATATTTTATGCTGTAAGTTTCTTCTGCTGCTGCTTTTGGCTTATTTGGTATGTTCCAT 4353 Db  
3457 CAGTTATTTCTGTTCACTTATTTGCTTCTTCCCTTTAAAGAAATTTTAAATATCAAAAG 3516 Qy  
4354 CAGTTATTTCTGTTCACTTATTTGCTTCTTCCCTTTAAAGAAATTTTAAATATCAAAAG 4413 Db  
3517 AATTTTGGTCAATTTATCTTCTGCTGCTGCTGCTTGTATTCGAATCACTGGAAGATGAC 3576 Qy  
4414 AATTTTGGTCAATTTATCTTCTGCTGCTGCTGCTTGTATTCGAATCACTGGAAGATGAC 4473 Db  
3577 CTTTCTTTATGGGATACACAAATTTGCAACTATTTCTTCTTATGCTTATGCTTATGCTATTC 3636 Qy  
4474 CTTTCTTTATGGGATACACAAATTTGCAACTATTTCTTCTTATGCTTATGCTATTCATTC 4533 Db  
3637 CAATCTATCCACTTCTAGTTGCTGATTTCTTCTTCAATAAGATTTCTTGGAAAGATGTAC 3696 Qy  
4534 CAATCTATCCACTTCTAGTTGCTGATTTCTTCTTCAATAAGATTTCTTGGAAAGATGTAC 4593 Db  
3697 GAAAAATGTGACACCTTATATCCATGGATAGGCTTTTCACTAGCTGTTATATGCTTATGCT 3756 Qy  
4594 GAAAAATGTGACACCTTATATCCATGGATAGGCTTTTCACTAGCTGTTATATGCTTATGCT 4653 Db  
3757 ACTGCACTGCTGCTGCTGCTTCTTCTTCAATACTATGAGAAATAATATGGAGGCA 3816 Qy  
4654 ACTGCACTGCTGCTGCTGCTTCTTCTTCAATACTATGAGAAATAATATGGAGGCA 4713 Db  
3817 GATCAATAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGCTTAAAAATAGGAAGC 3876 Qy  
4714 GATCAATAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGCTTAAAAATAGGAAGC 4773 Db  
3877 TTTCCAGAACCCAGCAATATGAGGATGAAGATGAAGTCAAGCTGAAAAAGCTTAAAGG 3936 Qy  
4774 TTTCCAGAACCCAGCAATATGAGGATGAAGATGAAGTCAAGCTGAAAAAGCTTAAAGG 4833 Db  
3937 TCAAGAGCTGATGGTGGCTGCTGCTGCTTGTGAGGAGAAACCATTCATTTATGCTCAGCAAT 3996 Qy  
4834 TCAAGAGCTGATGGTGGCTGCTGCTGCTTGTGAGGAGAAACCATTCATTTATGCTCAGCAAT 4893 Db  
3997 TGCATTAAGAAATATGATGACAAAGAAATTTTCTTCTTTTCAAGAAAGTAAAGAGTGG 4056 Qy  
4894 TGCATTAAGAAATATGATGACAAAGAAATTTTCTTCTTTTCAAGAAAGTAAAGAGTGG 4953 Db

QY 4057 CAACATAACATCTCTTTCTGTCGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAA 4116  
DB 4954 CAACATAACATCTCTTTCTGTCGAAAAAGGAGAGATCTTAGGACTATTGGGTCCAA 5013  
QY 4117 ATGGTCTGGCAAAACACAAATTAATAATTTCTGTTGGTGATATTGAACCACTTCAG 4176  
DB 5014 ATGGTCTGGCAAAACACAAATTAATAATTTCTGTTGGTGATATTGAACCACTTCAG 5073  
QY 4177 GCCAGTATTTTAGAGAGATATTCTTCAGAGACAAAGTGAAGATGATGATCACTGAAGT 4236  
DB 5074 GCCAGTATTTTAGAGAGATATTCTTCAGAGACAAAGTGAAGATGATGATCACTGAAGT 5133  
QY 4237 GTATGGTTACTGCTCTCAGATAAACCCTTTCTGCGCAGATACACTTCAGAGCAACTT 4296  
DB 5134 GTATGGTTACTGCTCTCAGATAAACCCTTTCTGCGCAGATACACTTCAGAGCAACTT 5193  
QY 4297 TTGAATTTATGGAGCTGTCAAAGGAATGAGTGCAAGTGACATGAAGAAGTCAATAGTC 4356  
DB 5194 TTGAATTTATGGAGCTGTCAAAGGAATGAGTGCAAGTGACATGAAGAAGTCAATAGTC 5253  
QY 4357 GAATACACATCATCTGATTTAAAGACATCTTCAGAGACTGTAAAGAACTACCTG 4416  
DB 5254 GAATACACATCATCTGATTTAAAGACATCTTCAGAGACTGTAAAGAACTACCTG 5313  
QY 4417 CAGGAATCAACGAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTT 4476  
DB 5314 CAGGAATCAACGAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTT 5373  
QY 4477 TCGTAGAACATCTCAGATGTGATCCCAAGCCAAACAGCACATGTGGCGAGCAA 4536  
DB 5374 TCGTAGAACATCTCAGATGTGATCCCAAGCCAAACAGCACATGTGGCGAGCAA 5433  
QY 4537 TTCCGAATCAATTTAAAGCAAGCGGCTGCTATTCTGACCACTCAGCTATATGAGG 4596  
DB 5434 TTCCGAATCAATTTAAAGCAAGCGGCTGCTATTCTGACCACTCAGCTATATGAGG 5493  
QY 4597 AGCGAGAGCTGCTGTGATCGATGATGATGATGATGATGATGATGATGATGATGATG 4656  
DB 5494 AGCGAGAGCTGCTGTGATCGATGATGATGATGATGATGATGATGATGATGATGATG 5553  
QY 4657 GAACAGTACAACATCTAAGAGTAAATTTGGAAAGGCTCTTTTGGAAATTAATGA 4716  
DB 5554 GAACAGTACAACATCTAAGAGTAAATTTGGAAAGGCTCTTTTGGAAATTAATGA 5613  
QY 4717 AGGACTGGATAGAAACCTTAGAGTAGACCCCTTTCAAAGAGAAATTCAGTATATTTCC 4776  
DB 5614 AGGACTGGATAGAAACCTTAGAGTAGACCCCTTTCAAAGAGAAATTCAGTATATTTCC 5673  
QY 4777 CAATGCAAGCGCTCAGGAAGTTTCTCTATTTTGGCTTATAAATTCCTAAGGAG 4836  
DB 5674 CAATGCAAGCGCTCAGGAAGTTTCTCTATTTTGGCTTATAAATTCCTAAGGAG 5733  
QY 4837 ATGTTCAAGTCCCTTTCAAACTCTTTTAAAGTGGAAAGCTAAACATGCTTTTGCCA 4896  
DB 5734 ATGTTCAAGTCCCTTTCAAACTCTTTTAAAGTGGAAAGCTAAACATGCTTTTGCCA 5793  
QY 4897 TTGAGAAATATAGCTTTTCTCAAGCAATTTGGAACAGGTTTGTGAGAACTCACTAAG 4956  
DB 5794 TTGAGAAATATAGCTTTTCTCAAGCAATTTGGAACAGGTTTGTGAGAACTCACTAAG 5853  
QY 4957 AACAGAGAGGAGGAGATATAGTTGGAACCTTTTAAACAGCACACTTTGGTGGGAACGAA 5016  
DB 5854 AACAGAGAGGAGGAGATATAGTTGGAACCTTTTAAACAGCACACTTTGGTGGGAACGAA 5913  
QY 5017 CACAAGAGATAGATAGTATTTGAATTTGTATGTTGGTCTGCTGCTACTGGGACTTCT 5076  
DB 5914 CACAAGAGATAGATAGTATTTGAATTTGTATGTTGGTCTGCTGCTACTGGGACTTCT 5973  
QY 5077 TTTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGTTTATTTTATGGAAGTGAACCTGGA 5136  
DB 5974 TTTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGTTTATTTTATGGAAGTGAACCTGGA 6033  
QY 5137 GAACCAAGACGCACTTGAATTTTCTTAAGCTCCTTTAATTTGAATCTGCTGTTGTG 5196

DB 6034 GAACCAAGACGCACTTGAATTTTCTTAAGCTCCTTAATTTGAATCTGCTGTTGTG 6093  
QY 5197 TTTTGTCTTTCTTTAAATAAAGTATGATATAATTAAGTGAAGTGCATGTTGTATTGA 5256  
DB 6094 TTTTGTCTTTCTTTAAATAAAGTATGATATAATTAAGTGAAGTGCATGTTGTATTGA 6153  
QY 5257 AGTATATTTGAACATATATAGTTTGTATGTCATCTTTTTCACCATTCAGAACAGTGTCT 5316  
DB 6154 AGTATATTTGAACATATATAGTTTGTATGTCATCTTTTTCACCATTCAGAACAGTGTCT 6213  
QY 5317 GAATTTGTGATTTAAAGGAATTTGTAATAGTAATGTTTATTTTAAAGTATCTTTAAGT 5376  
DB 6214 GAATTTGTGATTTAAAGGAATTTGTAATAGTAATGTTTATTTTAAAGTATCTTTAAGT 6273  
QY 5377 TATGCCATCTTTTAAATAAGTACGTAATGTTTCCAAATCTAAATAAAAACTAAT 5430  
DB 6274 TATGCCATCTTTTAAATAAGTACGTAATGTTTCCAAATCTAAATAAAAACTAAT 6327

RESULT 3  
AAD33648  
ID AAD33648 standard; cdna; 6369 BP.  
XX  
AC AAD33648;  
XX  
DT 01-JUL-2002 (first entry)  
XX  
DE Human TRICH-3 cdna.  
XX  
KW Human; transporter and ion channel; TRICH-3; transport disorder; angina;  
KW amyotrophic lateral sclerosis; cystic fibrosis; neuromuscular disorder;  
KW cardiac disorder; polymyositis; diabetes; neurological disorder; cancer;  
KW depression; schizophrenia; anaemia; Wilson's disease; Cushing's disease;  
KW cell proliferated disorder; infertility; arteriosclerosis; gene therapy;  
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; allergy;  
KW myasthenia gravis; multiple sclerosis; metabolic disorder; hypertension;  
KW acquired immune deficiency syndrome; immunological disorder; scleroderma;  
KW endocrine disorder; autoimmune thyroiditis; rheumatoid arthritis; goitre;  
KW cardiac myopathy; amnesia; toxic myopathy; Addison's disease; infection;  
KW epilepsy; mental disorder; myocarditis; Crohn's disease; Grave's disease;  
KW muscle disorder; stroke; dementia; anxiety; AIDS; asthma; cirrhosis;  
KW gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS  
FT 1238..6166  
FT /cag= a  
FT /product= "Human TRICH-3 protein"  
XX  
PN WO200212340-A2.  
XX  
XX 14-FEB-2002.  
XX  
XX 01-AUG-2001; 2001WO-US024217.  
XX  
XX 03-AUG-2000; 2000US-0223269P.  
XX 10-AUG-2000; 2000US-0224456P.  
XX 18-AUG-2000; 2000US-0226410P.  
XX 25-AUG-2000; 2000US-0228140P.  
XX 31-AUG-2000; 2000US-0230067P.  
XX 08-SEP-2000; 2000US-0231434P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Yue H, Thornton M, Ramkumar J, Tang YF, Azimzai Y, Baughn MR;  
XX Yang J, Yao MG, Lal P, Wallia NK, Gandhi AR, Hafalia AJA, Nguyen DB;  
XX Patterson C, Elliott VS, Tribouley CM, Lu DAM, Xu Y, Reddy R;  
XX Hernandez R, Borowsky ML, Lo TP, Lu Y, Policky JL, Greene BD;  
XX Sanjanwala MS, Raumann BE, Burford N, Ison CH, Lee EA, Ding L;  
XX Das D, Kallick DA, Khan FA, Seilhamer JJ;



1537 AAGAGCCATAGAAATTAGTGTATTCAGAGACATACAGAAAGAGGTGAAATGTGG 1596  
 2661 AAGAAGCCATAGAAATTAGTGTATTCAGAGACATACAGAAAGAGGTGAAATGTGG 2720  
 1597 AGGCTTTCAGAAATTTGTCATTTGACATATATGAGGTCAGATTACTGCTTACTTGCC 1656  
 2721 AGGCTTTCAGAAATTTGTCATTTGACATATATGAGGTCAGATTACTGCTTACTTGCC 2780  
 1657 ACAGTGGAAACAGAAAGAGTACATTTGATGAATATTTCTTTGCGACTCTGCCACCTTCTG 1716  
 2781 ACAGTGGAAACAGAAAGAGTACATTTGATGAATATTTCTTTGCGACTCTGCCACCTTCTG 2840  
 1717 ATGGCTTTGCTATATATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 1776  
 2841 ATGGGTTTGCATCTATATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 2900  
 1777 GAAATGATTTGGCAATTTGTCACAGTTAGATATACATTTTGATGTTTGGACAGTAGAAG 1836  
 2901 GAAATGATTTGGCAATTTGTCACAGTTAGATATACATTTTGATGTTTGGACAGTAGAAG 2960  
 1837 AAAATTTATCAATTTTGGCTTCAATCAAAAGGATACAGCCCAATATATACAGAAAG 1896  
 2961 AAAATTTATCAATTTTGGCTTCAATCAAAAGGATACAGCCCAATATATACAGAAAG 3020  
 1897 TGCAGAGGTTTACTAGATTTAGACATGCGAGCTATCAAAAGATAACCAAGCTTAAAAAT 1956  
 3021 TGCAGAGGTTTACTAGATTTAGACATGCGAGCTATCAAAAGATAACCAAGCTTAAAAAT 3080  
 1957 TAAGTGGTGTCAAAAAAGAAAGCTGTCATTTAGGAATTTGCTTTTGGGAACCCAAAGA 2016  
 3081 TAAGTGGTGTCAAAAAAGAAAGCTGTCATTTAGGAATTTGCTTTTGGGAACCCAAAGA 3140  
 2017 TACTGCTGTAGATCAACAGCTGGAATGGAGCCCTGTTCTCGACATATTTGTATGGA 2076  
 3141 TACTGCTGTAGATCAACAGCTGGAATGGAGCCCTGTTCTCGACATATTTGTATGGA 3200  
 2077 ATCTTTTAAATPACAGAAAGCCAAATCGGTTGACAGTGTTCAGTACTCATTTCTATGATG 2136  
 3201 ATCTTTTAAATPACAGAAAGCCAAATCGGTTGACAGTGTTCAGTACTCATTTCTATGATG 3260  
 2137 AGCTGACATTTCTCAGATAGGAAGCTGTGATATCAAGAGATGCTGGAATGTCTTG 2196  
 3261 AAGCTGACATTTCTCAGATAGGAAGCTGTGATATCAAGAGATGCTGGAATGTCTTG 3320  
 2197 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
 3321 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3380  
 2257 ACAATATTTGCCACAGAACTCTTCTTCTCAGTGTAAACACATATACCTCGAGCTA 2316  
 3381 ACAATATTTGCCACAGAACTCTTCTTCTCAGTGTAAACACATATACCTCGAGCTA 3440  
 2317 CTTTATTACAGAGATGACCAACAACTTTGTATAGTCTGCTTTCAAGGACATGGACA 2376  
 3441 CTTTATTACAGAGATGACCAACAACTTTGTATAGTCTGCTTTCAAGGACATGGACA 3500  
 2377 AATTTTCAGGTTTGTCTGCTTGCCTAGACAGTCAATCAAAATTTGGTGTCAATTTCTATG 2436  
 3501 AATTTTCAGGTTTGTCTGCTTGCCTAGACAGTCAATCAAAATTTGGTGTCAATTTCTATG 3560  
 2437 GTGTTTCCATGACAGCTTTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAAAATG 2496  
 3561 GTGTTTCCATGACAGCTTTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAAAATG 3620  
 2497 ACCAAGCAGATTTAGTGTATTTACTCAGCAGCCACTGGAGGAAAGAAATGGATTCAAAAT 2556  
 3621 ACCAAGCAGATTTAGTGTATTTACTCAGCAGCCACTGGAGGAAAGAAATGGATTCAAAAT 3680  
 2557 CTTTGTATCAATGGAACAGAGCTTACATTTCTTCTGGAACCAAGCTTCTCTAGTGA 2616  
 3681 CTTTGTATCAATGGAACAGAGCTTACATTTCTTCTGGAACCAAGCTTCTCTAGTGA 3740  
 2617 GCACCATGAGCCCTTTGGAAACAAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTACCT 2676

3741 GCACCATGAGCCCTTTGAAACCAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTACCT 3800  
 2677 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736  
 3801 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 3860  
 2737 TTCAAGATTTTATGTTTGTGTTTCATCACTCTTTTAAAAAATGCTGTGTTCCCATCAAAAC 2796  
 3861 TTCAAGATTTTATGTTTGTGTTTCATCACTCTTTTAAAAAATGCTGTGTTCCCATCAAAAC 3920  
 2797 TTGTTCCAGACTATATTTTCTAAACCTGGAGACAAACACATATAAACAACAAAGTC 2856  
 3921 TTGTTCCAGACTATATTTTCTAAACCTGGAGACAAACACATATAAACAACAAAGTC 3980  
 2857 TGCTTCTTCAAAATTTCTGCTGACTCAGATATCAGTGAATCTTATTAGCTTTTTCACAAGCC 2916  
 3981 TGCTTCTTCAAAATTTCTGCTGACTCAGATATCAGTGAATCTTATTAGCTTTTTCACAAGCC 4040  
 2917 AGAACAATATGTTGAGGATGATTAATGACAGTGAATGATATCCGTGGCTCCCATAGTG 2976  
 4041 AGAACAATATGTTGAGGATGATTAATGACAGTGAATGATATCCGTGGCTCCCATAGTG 4100  
 2977 CGCTTTTAAATGTGATGCTATTCAGAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTA 3036  
 4101 CGCTTTTAAATGTGATGCTATTCAGAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTA 4160  
 3037 CTATGTTTATTTCTTTTACCTATATTTAGTGAATATCTATAGTAACTACTATCTTATCAT 3096  
 4161 CTATGTTTATTTCTTTTACCTATATTTAGTGAATATCTATAGTAACTACTATCTTATCAT 4220  
 3097 TAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATTTTCAAGAAATTAAGTATA 3156  
 4221 TAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATTTTCAAGAAATTAAGTATA 4280  
 3157 TAGTTTTTAAAAATGAGCTGTATTTTCAAGAGCTTTGCTTGGAAATCATTTTCTGCAA 3216  
 4281 TAGTTTTTAAAAATGAGCTGTATTTTCAAGAGCTTTGCTTGGAAATCATTTTCTGCAA 4340  
 3217 TGCACCTTTACTTTTGCATGGAATAATGACAGAAATCAAGATCAAGCTTTATACCAAC 3276  
 4341 TGCACCTTTACTTTTGCATGGAATAATGACAGAAATCAAGATCAAGCTTTATACCAAC 4400  
 3277 TTAACCTTTCAAGTCTTTTGCATGCAATATGGAATGGAACAAGCTGTGTTGATATCC 3336  
 4401 TTAACCTTTCAAGTCTTTTGCATGCAATATGGAATGGAACAAGCTGTGTTGATATCC 4460  
 3337 CCTTATTTTATCAATTTTATTTTATGCTAGGAGCTTATTTGCAATTTCAATTTATGAT 3396  
 4461 CCTTATTTTATCAATTTTATTTTATGCTAGGAGCTTATTTGCAATTTCAATTTATGAT 4520  
 3397 TATATTTTATATCTGTAAAGTTCCTTTGCTGTGCTTTTTCCTTATTTGCTTATGTTCCAT 3456  
 4521 TATATTTTATATCTGTAAAGTTCCTTTGCTGTGCTTTTTCCTTATTTGCTTATGTTCCAT 4580  
 3457 CAGTATTTCTGTTCACTTATATGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAG 3516  
 4581 CAGTATTTCTGTTCACTTATATGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAG 4640  
 3517 AATTTTGGTCAATTTATCTATTTCTGCGCAGCTTTGCTTGTATTTGCAATCACTGAAATAA 3576  
 4641 AATTTTGGTCAATTTATCTATTTCTGCGCAGCTTTGCTTGTATTTGCAATCACTGAAATAA 4700  
 3577 CTTTCTTTATGGGATACAAATTTGCAATTTCTTCAATTTATGCTTTTGTATCATCATTC 3636  
 4701 CTTTCTTTATGGGATACAAATTTGCAATTTCTTCAATTTATGCTTTTGTATCATCATTC 4760  
 3637 CAATCTATCCACTTTCTAGGTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3696  
 4761 CAATCTATCCACTTTCTAGGTTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4820  
 3697 GAAAAATGTGGACACCTATAATCCATGGGATAGGCTTTTCAAGTGTATTTATGCTTTT 3756

4821 GAAGAAATGTCACACCTATAATCCATGGGATAGGCTTTCACTAGCTGTATATCGCCTT 4880  
4880  
3757 AACTCAGTGTCTACTGTGGATTTTCCCTTTCACTACTATCAGAAATAATCGAGCA 3816  
3816  
4881 AACTGAGTGTCTACTGTGGATTTTCCCTTTCACTACTATCAGAAATAATCGAGCA 4940  
4940  
3817 GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTAAATAAGGAGC 3876  
3876  
4941 GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTAAATAAGGAGC 5000  
5000  
3877 TTCCAGAACCCACAGACATGAGGATGAGATGAAGATGTCAGAGCTGAAGACTAAAGG 3936  
3936  
5001 TTCCAGAACCCACAGACATGAGGATGAGATGAAGATGTCAGAGCTGAAGACTAAAGG 5060  
5060  
3937 TCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCATCCATTATGCTCAGCAATT 3996  
3996  
5061 TCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCATCCATTATGCTCAGCAATT 5120  
5120  
3997 TGCATAAAGAATATGATGACGAAGAAGATTTTCTTTTCAAGAAAGTAAGAAAGTGG 4056  
4056  
5121 TGCATAAAGAATATGATGACGAAGAAGATTTTCTTTTCAAGAAAGTAAGAAAGTGG 5180  
5180  
4057 CAACTAAATACATCTCTTCTCTGTGTAAGAAAGAGAGATCTTAGGACTATTGGGTCCAA 4116  
4116  
5181 CAACTAAATACATCTCTTCTCTGTGTAAGAAAGAGAGATCTTAGGACTATTGGGTCCAA 5240  
5240  
4117 ATGGTGTGGGAAAGCAATATTAATAATTTCTGGTGGTGATATTGAACCAACTTCAG 4176  
4176  
5241 ATGGTGTGGGAAAGCAATATTAATAATTTCTGGTGGTGATATTGAACCAACTTCAG 5300  
5300  
4177 GCGAGTATTTTACGAGATTTCTTCCAGAGACAGTGAAGATGATGATTCATCTGAAGT 4236  
4236  
5301 GCGAGTATTTTACGAGATTTCTTCCAGAGACAGTGAAGATGATGATTCATCTGAAGT 5360  
5360  
4237 GTATGGGTTACTGCTCCTCAGATAAACCTTTGTGGCCAGATACACTATGCGAGAACATT 4296  
4296  
5361 GTATGGGTTACTGCTCCTCAGATAAACCTTTGTGGCCAGATACACTATGCGAGAACATT 5420  
5420  
4297 TTGAATTTATGGAGCTGTCAAGGAATGAGTGCAGTGAAGTGAAGTGAAGTGAAGTGA 4356  
4356  
5421 TTGAATTTATGGAGCTGTCAAGGAATGAGTGCAGTGAAGTGAAGTGAAGTGAAGTGA 5480  
5480  
4357 GAATAACACATCCTTGATTTTAAAGAACATCTTTTCAAGAACCTGTAAAGAAATCACTG 4416  
4416  
5481 GAATAACACATCCTTGATTTTAAAGAACATCTTTTCAAGAACCTGTAAAGAAATCACTG 5540  
5540  
4417 CAGGAATCAAGAAAGTGTGTTTCTCTAAGTATCTAGGAACTCTCAGATTAATT 4476  
4476  
5541 CAGGAATCAAGAAAGTGTGTTTCTCTAAGTATCTAGGAACTCTCAGATTAATT 5600  
5600  
4477 TGCTAGATGAACCATCTACAGTATGGATCCCAAGCCAAACAGACATGTGGCGAGCAA 4536  
4536  
5601 TGCTAGATGAACCATCTACAGTATGGATCCCAAGCCAAACAGACATGTGGCGAGCAA 5660  
5660  
4537 TTGCAACTGCATTTTAAACAGAAAGCGGCTGCTATTCTGACCACTCCTATATGAGG 4596  
4596  
5661 TTGCAACTGCATTTTAAACAGAAAGCGGCTGCTATTCTGACCACTCCTATATGAGG 5720  
5720  
4597 AGCGAGAGCTGCTGTGATCGAGTAGTATCATCTGCTGTGGCAGTTAAGATGATCG 4656  
4656  
5721 AGCGAGAGCTGCTGTGATCGAGTAGTATCATCTGCTGTGGCAGTTAAGATGATCG 5780  
5780  
4657 GAACAGTCAACATCTAAAGAGTAAATTTGAAAGAGGCTATTTTGGAAATTTAAATGA 4716  
4716  
5781 GAACAGTCAACATCTAAAGAGTAAATTTGAAAGAGGCTATTTTGGAAATTTAAATGA 5840  
5840  
4717 AGGCTGGATGAACCACTAGAGTAGACCCCTTCAAGAGAAATTCAGTATATTTCC 4776  
4776  
5841 AGGCTGGATGAACCACTAGAGTAGACCCCTTCAAGAGAAATTCAGTATATTTCC 5900  
5900  
4777 CAAATGCAAGCCGTCAGGAAAGTCTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAG 4836  
4836  
5901 CAAATGCAAGCCGTCAGGAAAGTCTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAG 5960  
5960

QY 4837 ATGTTTCAGTCCCTTTTCACTATCTTTTAACTGGAAGCACTTTTGGTGGGACGAA 4896  
4896  
Db 5961 ATGTTTCAGTCCCTTTTCACTATCTTTTAACTGGAAGCACTTTTGGTGGGACGAA 6020  
6020  
QY 4897 TTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAACAGCTTTTGTAGAACTCACTAAAG 4956  
4956  
Db 6021 TTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAACAGCTTTTGTAGAACTCACTAAAG 6080  
6080  
QY 4957 AACACAGAGGAGGAAGATATAGTTTGGAACTTTTAAACAGCACCTTTTGGTGGGACGAA 5016  
5016  
Db 6081 AACACAGAGGAGGAAGATATAGTTTGGAACTTTTAAACAGCACCTTTTGGTGGGACGAA 6140  
6140  
QY 5017 CACAAGAATATAGAGTAGTATTTTGAATTTGTTTGGTCTGCTTACTGGGACTTCT 5076  
5076  
Db 6141 CACAAGAATATAGAGTAGTATTTTGAATTTGTTTGGTCTGCTTACTGGGACTTCT 6200  
6200  
QY 5077 TTCTTTTTCACITTAATTTTAACTTTGGTTTAAAGTTTTATTGGAACTGTAACGCA 5136  
5136  
Db 6201 TTCTTTTTCACITTAATTTTAACTTTGGTTTAAAGTTTTATTGGAACTGTAACGCA 6260  
6260  
QY 5137 GAACCAAGAACGCACTTGAATTTTCTAAGCTCCTTAATTTGAAATGCTGTGTTGTGTG 5196  
5196  
Db 6261 GAACCAAGAACGCACTTGAATTTTCTAAGCTCCTTAATTTGAAATGCTGTGTTGTGTG 6320  
6320  
QY 5197 TTTTGTCTTTCTTTTAAATTAAGATGATGATTAATAA 5233  
5233  
Db 6321 TTTTGTCTTTCTTTTAAATTAAGATGATGATTAATAA 6357  
6357

## RESULT 4

AAD37620  
ID AAD37620 standard; cDNA; 5262 BP.

XX AAD37620;

AC AC  
XX XX  
DT 10-SEP-2002 (first entry)

XX Human transporter protein cDNA #3.

XX Human; novel human protein; NHP; transporter protein; mental disorder;  
KW cancer; gene therapy; drug screening; nutraceutical application;  
KW cosmetic application; polymorphism; ss.

XX Homo sapiens.

XX WO200231147-A2.

XX 18-APR-2002.

XX 04-OCT-2001; 2001WO-US031113.

XX 10-OCT-2000; 2000US-0239629P.

XX (LEXI-) LEXICON GENETICS INC.

XX Hu Y, Nepomnichy B;

XX WPI; 2002-454552/48.

XX Novel nucleic acid molecule encoding novel human proteins, useful for  
therapeutic, diagnostic and pharmacogenomic applications.

XX Disclosure; Page 45-46; 46pp; English.

XX The present sequence is a cDNA coding for novel human protein (NHP),  
CC human transporter protein. NHPs shares structural similarity with the  
CC mammalian ATP-binding cassette (ABC) transporters and multidrug  
CC resistance transporters. NHP polynucleotides are useful for the  
CC therapeutic, diagnostic and pharmacogenomic applications. They are used  
CC for detecting and treating mental disorders and cancers. They are also  
CC used in gene therapy. NHP polypeptides are useful for diagnosis, drug  
CC screening, clinical trial monitoring, treatment of diseases and



CC disorders, and cosmetic or nutraceutical applications

Sequence 5262 BP; 1615 A; 888 C; 1012 G; 1743 T; 0 U; 4 Other;

Query Match 85.2%; Score 4664; DB 6; Length 5262;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5014; Conservative 0; Mismatches 7; Indels 0

Qy	97	AGGTTTATT	CAGAAAA	CATGTC	CACTG	CAATATTAGG	AGGTAGG	AGTTTGG	A	CAGACCA	156
Db	238	AGGTTTATT	TTCAGAAAA	CATGTCC	CACTG	CAATATTAGG	AGGTAGG	AGTTTGG	A	CAGACCA	297
Qy	157	GAA	CAC	TCT	CACT	CA	GAGAA	TAC	TTAA	TTC	216
Db	298	GAA	CAC	TCT	CACT	CA	GAGAA	TAC	TTAA	TTC	357
Qy	217	AAAT	TCT	TTTT	TTCC	ACTAT	TTTTTTTT	TTATTT	TG	GT	276
Db	358	AAAT	TCT	TTTT	TTCC	ACTAT	TTTTTTTT	TTATTT	TG	GT	417
Qy	277	ATA	GAA	AAT	TAT	GAA	GAG	TG	CGT	CT	336
Db	418	ATA	GAA	AAT	TAT	GAA	GAG	TG	CGT	CT	477
Qy	337	CTA	AT	CT	AA	TCT	TTGG	A	TAT	ACT	396
Db	478	CTA	AT	CT	AA	TCT	TTGG	A	TAT	ACT	537
Qy	397	TGT	CT	ACT	GA	T	CA	T	CA	T	456
Db	538	TGT	CT	ACT	GA	T	CA	T	CA	T	597
Qy	457	TG	T	TA	A	C	A	C	A	G	516
Db	598	TG	T	TA	A	C	A	C	A	G	657
Qy	517	TGT	C	T	A	T	G	A	A	C	576
Db	658	TGT	C	T	A	T	G	A	A	C	717
Qy	577	CA	G	A	G	C	T	G	G	T	636
Db	718	CA	G	A	G	C	T	G	G	T	777
Qy	637	TTTT	T	A	C	A	A	G	C	A	696
Db	778	TTTT	T	A	C	A	A	G	C	A	837
Qy	697	AGG	A	G	C	T	G	G	A	A	756
Db	838	AGG	A	G	C	T	G	G	A	A	897
Qy	757	C	T	T	T	C	C	C	C	C	816
Db	898	C	T	T	T	C	C	C	C	C	957
Qy	817	TTTT	T	G	G	C	A	A	T	T	876
Db	958	TTTT	T	G	G	C	A	A	T	T	1017
Qy	877	T	G	G	A	C	T	T	T	T	936
Db	1018	T	G	G	A	C	T	T	T	T	1077
Qy	937	T	T	C	T	A	T	G	T	C	996
Db	1078	T	T	C	T	A	T	G	T	C	1137
Qy	997	G	C	A	C	A	T	T	T	T	1056
Db	1138	G	C	A	C	A	T	T	T	T	1197
Qy	1057	T	A	A	T	T	T	T	T	T	1116



Db 2278 ANGTCGACATCTTCGACATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTGTG 2337  
Qy 2197 GTTCTTCAATGTTCTCAAAAGTAATGGGGATCGGCTACGGCTGAGCANGTACATAG 2256  
Db 2338 GTTCTTCAATGTTCTCAAAAGTAATGGGGATCGGCTACGGCTGAGCANGTACATAG 2397  
Qy 2257 ACAAAATATTGCGCCACAGAACTCTTTCTTCACTGGTTTAAACAACATATACCTGGAGCTA 2316  
Db 2398 ACAAAATATTGCGCCACAGAACTCTTTCTTCACTGGTTTAAACAACATATACCTGGAGCTA 2457  
Qy 2317 CTTTATTAACAAGAAATGCAACAACTTTGTTGTTATAGCTTGCCCTTCAAGGACATGACA 2376  
Db 2458 CTTTATTAACAAGAAATGCAACAACTTTGTTGTTATAGCTTGCCCTTCAAGGACATGACA 2517  
Qy 2377 AATTTTCAGGTTTGTGTTCTGCGCTAGACAGTCAATCAAAATTTGGGTGCTATTTCTTATG 2436  
Db 2518 AATTTTCAGGTTTGTGTTCTGCGCTAGACAGTCAATCAAAATTTGGGTGCGATTTCTTATG 2577  
Qy 2437 GTGTTTCCATGACGACTTTGGAAGACGTATTTTAAAGCTAGAAAGTTGAACAGAAATG 2496  
Db 2578 GGGTTTCCATGACGACTTTGGAAGACGTATTTTAAAGCTAGAAAGTTGAACAGAAATG 2637  
Qy 2497 ACCAAGCAGATTATAGTGTATTTACTCAGCAGCCACTGGAGGAAGAAATGATTCAAAAT 2556  
Db 2638 ACCAAGCAGATTATAGTGTATTTACTCAGCAGCCACTGGAGGAAGAAATGATTCAAAAT 2697  
Qy 2557 CTTTGTGAAATGGAACAGAGCTTACTTATTTCTTGAACCAAGCTTCTCTAGTGA 2616  
Db 2698 CTTTGTGAAATGGAACAGAGCTTACTTATTTCTTGAACCAAGCTTCTCTAGTGA 2757  
Qy 2617 GCACCATGAGCCTTTTGGAAACACAGATGTATACATAGCAAGTTTCATTTCTTTACCT 2676  
Db 2758 GCACCATGAGCCTTTTGGAAACACAGATGTATACATAGCAAGTTTCATTTCTTTACCT 2817  
Qy 2677 TGAACCGTGAAGTAATTAACAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG 2736  
Db 2818 TGAACCGTGAAGTAATTAACAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG 2877  
Qy 2737 TTCAGATTTTATGTTTTTGTGTTCACTCACTCTTTTAAAAATGCTGTGTTCCCATCAAC 2796  
Db 2878 TTCAGATTTTATGTTTTTGTGTTCACTCACTCTTTTAAAAATGCTGTGTTCCCATCAAC 2937  
Qy 2797 TTGTTTCCAGACTTATTTTCTTAAACCTGGAGACAAACCAATATAACAAACAAAGTC 2856  
Db 2938 TTGTTTCCAGACTTATTTTCTTAAACCTGGAGACAAACCAATATAACAAACAAAGTC 2997  
Qy 2857 TGCCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 2916  
Db 2998 TGCCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 3057  
Qy 2917 AGAACATATGTTGACGATGATTATGACAGTGACTATGATCCGTGCTCCCATAGTG 2976  
Db 3058 AGAACATATGTTGACGATGATTATGACAGTGACTATGATCCGTGCTCCCATAGTG 3117  
Qy 2977 CGGCTTTAAATGTGATGCAANTCAGAAAAGGACTATGTTTTTTCAGAGCTTTTTCACAGTA 3036  
Db 3118 CGGCTTTAAATGTGATGCAANTCAGAAAAGGACTATGTTTTTTCAGAGCTTTTTCACAGTA 3177  
Qy 3037 CTATGGTTTATCTTTTACCTATATTAGTCAATATCATTAGTAACCTACTATCTTTTATCAT 3096  
Db 3178 CTATGGTTTATCTTTTACCTATATTAGTCAATATCATTAGTAACCTACTATCTTTTATCAT 3237  
Qy 3097 TAAATGTGACTGAAACCACTCCAGATCTGGAGTACCCCAATCTTTTCAAGAAATTTACTGATA 3156  
Db 3238 TAAATGTGACTGAAACCACTCCAGATCTGGAGTACCCCAATCTTTTCAAGAAATTTACTGATA 3297  
Qy 3157 TAGTTTTTAAATGAGCTGATTTTTTCAAGAGCTTTTCTTGGATCATCTGTTACTGCAA 3216  
Db 3298 TAGTTTTTAAATGAGCTGATTTTTTCAAGAGCTTTTCTTGGATCATCTGTTACTGCAA 3357  
Qy 3217 TGCACCTTACTTTTGCATGGAATGCAAGAAATCATAAGATCAAAAGCTTATACTCAAC 3276  
Db 3358 TGCACCTTACTTTTGCATGGAATGCAAGAAATCATAAGATCAAAAGCTTATACTCAAC 3417

Qy 3277 TTAACCTTTCCAGTCTTTTGGCCATCTGCAATATGATGGACAAGCTGTGTTGATATCC 3336  
Db 3418 TTAACCTTTCCAGTCTTTTGGCCATCTGCAATATGATGGACAAGCTGTGTTGATATCC 3477  
Qy 3337 CTTATTTTTTTATCATCTTATTTTTGATGCTAGGAAGCTTATTTGGCATTTCTATTGAT 3396  
Db 3478 CTTATTTTTTTATCATCTTATTTTTGATGCTAGGAAGCTTATTTGGCATTTCTATTGAT 3537  
Qy 3397 TATATTTTTTATCTGTAAAGTTTCTTGTGCTGTGGTTTTTGGCTTATTTGGTTATGTTCCAT 3456  
Db 3538 TATATTTTTTATCTGTAAAGTTTCTTGTGCTGTGGTTTTTGGCTTATTTGGTTATGTTCCAT 3597  
Qy 3457 CAGTATTCTGTTCACCTATATTGCTTTCTTCACTTTAAAGAAAAATTTTAAATATCCAAAG 3516  
Db 3598 CAGTATTCTGTTCACCTATATTGCTTTCTTCACTTTAAAGAAAAATTTTAAATATCCAAAG 3657  
Qy 3517 AATTTTGGTCAATTAATCTATTTCTGTGGCAGCGTTGGCTGTGTTATTCGAATCATCTGAATTA 3576  
Db 3658 AATTTTGGTCAATTAATCTATTTCTGTGGCAGCGTTGGCTGTGTTATTCGAATCATCTGAATTA 3717  
Qy 3577 CTTTCTTTATGGGATACAAATTTGCAACTATTTCTTTCAATATGCTTTTGTATCATCATTC 3636  
Db 3718 CTTTCTTTATGGGATACAAATTTGCAACTATTTCTTTCAATATGCTTTTGTATCATCATTC 3777  
Qy 3637 CAATCTATCCACTTCTAGTTGCTGATTTCTTTCATAAAGATTTCTTGGAGAATGTAC 3696  
Db 3778 CAATCTATCCACTTCTAGTTGCTGATTTCTTTCATAAAGATTTCTTGGAGAATGTAC 3837  
Qy 3697 GAAAAATGTGACACACTATAATCCATGGATAGGCTTTTCACTAGCTGTATATGCGCTT 3756  
Db 3838 GAAAAATGTGACACACTATAATCCATGGATAGGCTTTTCACTAGCTGTATATGCGCTT 3897  
Qy 3757 ACCTGAGTGTGTACTGTGGATTTTCTCTTACATACTATGAGAAAAAATATGAGGCA 3816  
Db 3898 ACCTGAGTGTGTACTGTGGATTTTCTCTTACATACTATGAGAAAAAATATGAGGCA 3957  
Qy 3817 GATCAATAAGAAAAAGATCCCTTTTTCAGAAAACTTTTCAACGAAGCTTAAAAATAGGAAGC 3876  
Db 3958 GATCAATAAGAAAAAGATCCCTTTTTCAGAAAACTTTTCAACGAAGCTTAAAAATAGGAAGC 4017  
Qy 3877 TTCAGAACCCACAGCAATAGAGATGAAGATGAAGATGTCAAAGCTGAAAGCTAAAGG 3936  
Db 4018 TTCAGAACCCACAGCAATAGAGATGAAGATGAAGATGTCAAAGCTGAAAGCTAAAGG 4077  
Qy 3937 TCAGAGAGCTGATGGGTTGCCAGTGTGTGAGGGAACCATTCATTTATGCTCAGCAAT 3996  
Db 4078 TCAGAGAGCTGATGGGTTGCCAGTGTGTGAGGGAACCATTCATTTATGCTCAGCAAT 4137  
Qy 3997 TGCATAAAGAAATATGATGACAAGAAAGATTTTCTTTTCAAGAAAAAGTAAAGAGTGG 4056  
Db 4138 TGCATAAAGAAATATGATGACAAGAAAGATTTTCTTTTCAAGAAAAAGTAAAGAGTGG 4197  
Qy 4057 CAACTAAATACATCTCTTTTCTGTGAAAAAGGAGAGATCTTTCAGACTATTTGGGTCCAA 4116  
Db 4198 CAACTAAATACATCTCTTTTCTGTGAAAAAGGAGAGATCTTTCAGACTATTTGGGTCCAA 4257  
Qy 4117 ATGGTCTCGCAAAAGCACAAATTAATTAATTTCTGTTGGTGTATTTGAACCAACTTCAG 4176  
Db 4258 ATGGTCTCGCAAAAGCACAAATTAATTAATTTCTGTTGGTGTATTTGAACCAACTTCAG 4317  
Qy 4177 GCCAGGTAATTTTGGAGATTAATTTCTTTCAGAGCAAGTGAAGATGATGATCTCAGTGAAGT 4236  
Db 4318 GCCAGGTAATTTTGGAGATTAATTTCTTTCAGAGCAAGTGAAGATGATGATCTCAGTGAAGT 4377  
Qy 4237 GTATGGTTTACTGTCTCCTCAGATAAACCCCTTTTGTGGCCAGATACTACATTTGCGAGGAACAT 4296  
Db 4378 GTATGGTTTACTGTCTCCTCAGATAAACCCCTTTTGTGGCCAGATACTACATTTGCGAGGAACAT 4437  
Qy 4297 TTGAAATTTATGAGCTGTCAAAGGAATGAGTGCAGATGCAATGAAAGATCATAGTC 4356  
Db 4438 TTGAAATTTATGAGCTGTCAAAGGAATGAGTGCAGATGCAATGAAAGATCATAGTC 4497



Db 301 CTGATGCTATATTTACTGAGAAATATACAAATGAAAAGAAATGTTAAACATCCAGTCTC 360  
QY 474 TCTAAGCCGAGCAACTTTTGTAGGTGGTTTCAAAGACTCCATGCTCTATGAACCTTGGT 533  
Db 361 TCTAAGCCGAGCAACTTTTGTAGGTGGTTTCAAAGACTCCATGCTCTATGAACCTTGGT 420  
QY 534 TTTTTCCTGATGATGATCCAGTATCTCTATTTATATAGGATTCAGAGCTGGCTGTTCA 593  
Db 421 TTTTTCCTGATGATGATCCAGTATCTCTATTTATATAGGATTCAGAGCTGGCTGTTCA 480  
QY 594 AAATCATGTGAGGCTGCTCAGTATGCTGCTCAGGTTTCAAGTTTCAAGCATCCATA 653  
Db 481 AAATCATGTGAGGCTGCTCAGTATGCTGCTCAGGTTTCAAGTTTCAAGCATCCATA 540  
QY 654 GATGCTGCCATTATACAGTTGAGACCAATGTTCTCTTGAAGGAGCTGGAGTCACT 713  
Db 541 GATGCTGCCATTATACAGTTGAGACCAATGTTCTCTTGAAGGAGCTGGAGTCACT 600  
QY 714 AAAGCTGTTTATTATGGGAGAAACTGCTGTTGTAGAAATAGATPACCTTTCCCGGAGGTA 773  
Db 601 AAAGCTGTTTATTATGGGAGAAACTGCTGTTGTAGAAATAGATPACCTTTCCCGGAGGTA 660  
QY 774 ATTTTAATATACCTAGTATAGATTTTCACTTTTGGATCTTTTGGCAATTCATATC 833  
Db 661 ATTTTAATATACCTAGTATAGATTTTCACTTTTGGATCTTTTGGCAATTCATATC 720  
QY 834 GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATATGCGGACTTCATGATACT 893  
Db 721 GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATATGCGGACTTCATGATACT 780  
QY 894 GCCTTTTGGGTTTCCCTGGGTTCTCTATATACAGTTTAAATTTTCTATGTCCTTCTT 953  
Db 781 GCCTTTTGGGTTTCCCTGGGTTCTCTATATACAGTTTAAATTTTCTATGTCCTTCTT 840  
QY 954 ATGGCAGTCAATCGCAGACGCTCTTTGTTTATTTCTCTCAAGTAGCAGCATTTGTATATT 1013  
Db 841 ATGGCAGTCAATCGCAGACGCTCTTTGTTTATTTCTCTCAAGTAGCAGCATTTGTATATT 900  
QY 1014 CTGCTTTTCTTTTCTTTATGATATCACTGTATTTTGTCTTTTAAATGCTGACACTCTT 1073  
Db 901 CTGCTTTTCTTTTCTTTATGATATCACTGTATTTTGTCTTTTAAATGCTGACACTCTT 960  
QY 1074 TTTTAAAAATCAAAACATGTGGGATGTGATTTTGTGTTTGTACTGTGGCTTTGGATT 1133  
Db 961 TTTTAAAAATCAAAACATGTGGGATGTGATTTTGTGTTTGTACTGTGGCTTTGGATT 1020  
QY 1134 ATTTGGCTTATGATAATCCTCATAGAAAGTTTCCCAAATCGTTAGTGTGGCTTTTCAGT 1193  
Db 1021 ATTTGGCTTATGATAATCCTCATAGAAAGTTTCCCAAATCGTTAGTGTGGCTTTTCAGT 1080  
QY 1194 CCTTTCTGTACGTACTTTGTGATTTGTGATTTGACAGGTCATGCAATTTAGAGATTTT 1253  
Db 1081 CCTTTCTGTACGTACTTTGTGATTTGTGATTTGACAGGTCATGCAATTTAGAGATTTT 1140  
QY 1254 AATGAAGGTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCTCTAATTTATCAAT 1313  
Db 1141 AATGAAGGTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCTCTAATTTATCAAT 1200  
QY 1314 ATCATGCTCACACTTAATAGTATATTTCTATGTCCTTTGGCTGTCTATCTTGTATCAAGTC 1373  
Db 1201 ATCATGCTCACACTTAATAGTATATTTCTATGTCCTTTGGCTGTCTATCTTGTATCAAGTC 1260  
QY 1374 ATTTCCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGGCTTCATATTGG 1433  
Db 1261 ATTTCCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGGCTTCATATTGG 1320  
QY 1434 TCAAGAGCAAAAGAAATATAGAGGATTTATCAGAGGGCAATGTTAATGGAAATATTAGT 1493  
Db 1321 TCAAGAGCAAAAGAAATATAGAGGATTTATCAGAGGGCAATGTTAATGGAAATATTAGT 1380  
QY 1494 TTTTAGTGAATTTATCGAGCCAGTTTCTTCAGATTTGTAGGAAAGAGAGCCATAAGAAAT 1553

Db 1381 TTTTAGTGAATTTATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATTAAGAAAT 1440  
QY 1554 AGTGGTATTTCCAGAAAGCATACAGAAAGAGGTGAAAATGTGGAGGCTTTTGAAAATTTG 1613  
Db 1441 AGTGGTATTTCCAGAAAGCATACAGAAAGAGGTGAAAATGTGGAGGCTTTTGAAAATTTG 1500  
QY 1614 TCAATTTGACATATATCAGGCTCAGATTACTGCTTACTTTGGCCACACAGTGGAAACAGGAAAG 1673  
Db 1501 TCAATTTGACATATATCAGGCTCAGATTACTGCTTACTTTGGCCACACAGTGGAAACAGGAAAG 1560  
QY 1674 AGTACATTTGATGAAATTTCTTTGTGACCTCTGCCACCTTCTGTATGGGTTTGCATCTATA 1733  
Db 1561 AGTACATTTGATGAAATTTCTTTGTGACCTCTGCCACCTTCTGTATGGGTTTGCATCTATA 1620  
QY 1734 TATGGACACAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTGGCAAT 1793  
Db 1621 TATGGACACAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTGGCAAT 1680  
QY 1794 TGTCCACAGTTAGATATACACTTTTGTGATTTTGAAGTATAGAAATTTTCAATTTTGG 1853  
Db 1681 TGTCCACAGTTAGATATACACTTTTGTGATTTTGAAGTATAGAAATTTTCAATTTTGG 1740  
QY 1854 GCTTCAATCAAGGGGATACCAGCCACAATATATATCAAGAAAGTCAGAGGTTTACTA 1913  
Db 1741 GCTTCAATCAAGGGGATACCAGCCACAATATATATCAAGAAAGTCAGAGGTTTACTA 1800  
QY 1914 GATTAGACATGACAGCTATCAAGATAACCAAGCTTAAATAATTAAGTGGTGGTCAAAA 1973  
Db 1801 GATTAGACATGACAGCTATCAAGATAACCAAGCTTAAATAATTAAGTGGTGGTCAAAA 1860  
QY 1974 AGAAGCTGTCAATTAGGAATGCTGTTCTGGGAACCCAAAGATCTGCTGTAGATGAA 2033  
Db 1861 AGAAGCTGTCAATTAGGAATGCTGTTCTGGGAACCCAAAGATCTGCTGTAGATGAA 1920  
QY 2034 CCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGTATGGAATCTTTTAAATAACAGA 2093  
Db 1921 CCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGTATGGAATCTTTTAAATAACAGA 1980  
QY 2094 ABAAGCAATCGGCTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTTGA 2153  
Db 1981 ABAAGCAATCGGCTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTTGA 2040  
QY 2154 GATAGGAAGCTGTGATATCACAAGGAATGCTGAAAATGTTGGTTTCTTCAATTTGTTCTC 2213  
Db 2041 GATAGGAAGCTGTGATATCACAAGGAATGCTGAAAATGTTGGTTTCTTCAATTTGTTCTC 2100  
QY 2214 AAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATACACAAATATTTGTCACACA 2273  
Db 2101 AAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATACACAAATATTTGTCACACA 2160  
QY 2274 GAATCTCTTTCTTCTCCTGTTAAACACATATATACCTGGAGCTACTTTTATACACAGAAT 2333  
Db 2161 GAATCTCTTTCTTCTCCTGTTAAACACATATATACCTGGAGCTACTTTTATACACAGAAT 2220  
QY 2334 GACCAACACTGTGTATAGCTTGCCTTTCAAGGACATGAGCAAAATTTTTCAGGTTTGT 2393  
Db 2221 GACCAACACTGTGTATAGCTTGCCTTTCAAGGACATGAGCAAAATTTTTCAGGTTTGT 2280  
QY 2394 TCTGCCCCTAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTATGGTGTCTTCCATGAGCACT 2453  
Db 2281 TCTGCCCCTAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTATGGGTCTTCCATGAGCACT 2340  
QY 2454 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATTAGT 2513  
Db 2341 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATTAGT 2400  
QY 2514 GTATTTTACTCAGAGCCACTGGAGGAGAAATGGATTCAAAATCTTTTGTATGAAATGGAA 2573  
Db 2401 GTATTTTACTCAGAGCCACTGGAGGAGAAATGGATTCAAAATCTTTTGTATGAAATGGAA 2460  
QY 2574 CAGAGCTTACTTATTTCTGAAAACCAAGGCTTCTCTAGTAGCAGCCATGAGCCCTTGG 2633  
Db 2461 CAGAGCTTACTTATTTCTTCTGAAAACCAAGGCTKCTCTAGTAGCAGCCATGAGCCCTTGG 2520

QY 2634 AAAACAAGATGTATACAAATAGCAAGTTTCATTTCTTTTACCTTGAACGCTGAAGTAAA 2693  
DB 2521 AAAACAAGATGTATACAAATAGCAAGTTTCATTTCTTTTACCTTGAACGCTGAAGTAAA 2580  
QY 2694 TCAGTGAGATCAGTGTTCTCTCTGCTTTAAATTTTTTTTTCACAGTTCAGATTTTATGTTT 2753  
DB 2581 TCAGTGAGATCAGTGTTCTCTCTGCTTTAAATTTTTTTTTCACAGTTCAGATTTTATGTTT 2640  
QY 2754 TTGGTTTCATCTCTTTTAAAAATGCTGTGTTCCCATCAACCTTGTTCAGAGCTTATAT 2813  
DB 2641 TTGGTTTCATCTCTTTTAAAAATGCTGTGTTCCCATCAACCTTGTTCAGAGCTTATAT 2700  
QY 2814 TTTCTAAAAACCTGGAGACAAACCAATATAATACAAAAACAAGTCTGCTTCTTCAAAATTTCT 2873  
DB 2701 TTTCTAAAAACCTGGAGACAAACCAATATAATACAAAAACAAGTCTGCTTCTTCAAAATTTCT 2760  
QY 2874 GCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCGAGACAAATAGTGAAG 2933  
DB 2761 GCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCGAGACAAATAGTGAAG 2820  
QY 2934 ATGATTAAATGACAGTACTATGATCCGCTGCTCCCATAGTGGGCTTTAAATGTGATG 2993  
DB 2821 ATGATTAAATGACAGTACTATGATCCGCTGCTCCCATAGTGGGCTTTAAATGTGATG 2880  
QY 2994 CATTGAGAAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTACTATGTTTTTCTTTA 3053  
DB 2881 CATTGAGAAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTACTATGTTTTTCTTTA 2940  
QY 3054 CCTATATTAGTGAATATCATATTAGTAATCTATCTTTATCATTTAAATGTGACTGAAACC 3113  
DB 2941 CCTATATTAGTGAATATCATATTAGTAATCTATCTTTATCATTTAAATGTGACTGAAACC 3000  
QY 3114 ATCCAGATCTGGAGTACCCCACTCTTTTCAAGAAATTTACTGATATAGTTTTTAAATTTGAG 3173  
DB 3001 ATCCAGATCTGGAGTACCCCACTCTTTTCAAGAAATTTACTGATATAGTTTTTAAATTTGAG 3060  
QY 3174 CTGTATTTTCAAGAGCTTTGCTTTGGATCATTTGTTACTGCAATGCCACCTTACTTTGCC 3233  
DB 3061 CTGTATTTTCAAGAGCTTTGCTTTGGATCATTTGTTACTGCAATGCCACCTTACTTTGCC 3120  
QY 3234 ATGGAAAATGCAAGAAATCATAAGATCAAAGTCTATCTCAACTTAACTTTTCAGGTCTT 3293  
DB 3121 ATGGAAAATGCAAGAAATCATAAGATCAAAGTCTATCTCAACTTAACTTTTCAGGTCTT 3180  
QY 3294 TTGGCATCTGCATATGGAATGGACAAGCTGTTGTTGATATCCCTTATTTTTTATCATTT 3353  
DB 3181 TTGGCATCTGCATATGGAATGGACAAGCTGTTGTTGATATCCCTTATTTTTTATCATTT 3240  
QY 3354 CTTATTTTGTATGCTAGGAAGCTTTATTTGGCATTTTCAATTTATGGAATATATTTTTTACTGTA 3413  
DB 3241 CTTATTTTGTATGCTAGGAAGCTTTATTTGGCATTTTCAATTTATGGAATATATTTTTTACTGTA 3300  
QY 3414 AAGTTCTTGTGCTGTTTTTGGCTTATTTGGTTATGTTCCATCAGTTATTTCTGTTCACT 3473  
DB 3301 AAGTTCTTGTGCTGTTTTTGGCTTATTTGGTTATGTTCCATCAGTTATTTCTGTTCACT 3360  
QY 3474 TATATTGCTTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3533  
DB 3361 TATATTGCTTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
QY 3534 TATTCTGTGGAGCGTTGGCTTGTATTTGCAATCTAGTAAATTAATCTTTCTTTATGGGATAC 3593  
DB 3421 TATTCTGTGGAGCGTTGGCTTGTATTTGCAATCTAGTAAATTAATCTTTCTTTATGGGATAC 3480  
QY 3594 ACAATTTGCAACTATTTCTTATATGCTTTTGTATCATCTTCCATCTATCCACTTCTA 3653  
DB 3481 ACAATTTGCAACTATTTCTTATATGCTTTTGTATCATCTTCCATCTATCCACTTCTA 3540  
QY 3654 GGTGTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGACGAAAAAATTTGTGGACACC 3713  
DB 3541 GGTGTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGACGAAAAAATTTGTGGACACC 3600

QY 3714 TATAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATGCTTTTACCTGTCAGTGTGTACTG 3773  
DB 3601 TATAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATGCTTTTACCTGTCAGTGTGTACTG 3660  
QY 3774 TGGATTTTCTCTTACATATCTATGAGAAAAATATGAGGGCAGATCATATAGAAAGAT 3833  
DB 3661 TGGATTTTCTCTTACATATCTATGAGAAAAATATGAGGGCAGATCAATATAGAAAGAT 3720  
QY 3834 CCTTTTTCAGAAAACTTTTCAACGAAAGTCTAAAAATAGGAAGCTTCCAGAAACCAACAGAC 3893  
DB 3721 CCTTTTTCAGAAAACTTTTCAACGAAAGTCTAAAAATAGGAAGCTTCCAGAAACCAACAGAC 3780  
QY 3894 AATGAGGATGAGATGAGATGTCAAAGCTGAAAGCTTAAAGGTCAAAGAGCTGATGGGT 3953  
DB 3781 AATGAGGATGAGATGAGATGTCAAAGCTGAAAGCTTAAAGGTCAAAGAGCTGATGGGT 3840  
QY 3954 TGCAGTGTGTGAGGAGAAACCATCTATTATGCTGAGCAATTTGCTATGCAATGATATGAT 4013  
DB 3841 TGCAGTGTGTGAGGAGAAACCATCTATTATGCTGAGCAATTTGCTATGCAATGATATGAT 3900  
QY 4014 GACAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 4073  
DB 3901 GACAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 3960  
QY 4074 TTTCTGTGCAAAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGCTGCTGCAAAAGC 4133  
DB 3961 TTTCTGTGCAAAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGCTGCTGCAAAAGC 4020  
QY 4134 ACAATTTATTAATTTCTGTTGCTGATATTGAACCAACTTTCAGGCCAGGATATTTTATGGA 4193  
DB 4021 ACAATTTATTAATTTCTGTTGCTGATATTGAACCAACTTTCAGGCCAGGATATTTTATGGA 4080  
QY 4194 GATTATTTCTGAGAGCAAGTGAAGTGAATGATTTTCACTGAAGTGTATGGTTTACTGTCT 4253  
DB 4081 GATTATTTCTGAGAGCAAGTGAAGTGAATGATTTTCACTGAAGTGTATGGTTTACTGTCT 4140  
QY 4254 CAGATAAAACCTTTTGTGGCCAGATACATATTGTCAGGAAACATTTTGAATTTTATGAGCT 4313  
DB 4141 CAGATAAAACCTTTTGTGGCCAGATACATATTGTCAGGAAACATTTTGAATTTTATGAGCT 4200  
QY 4314 GTCAAAAGGAATGAGTGCACAGTGCATGAAGAAAGTCAATAGTTCGAATACATACATCT 4373  
DB 4201 GTCAAAAGGAATGAGTGCACAGTGCATGAAGAAAGTCAATAGTTCGAATACATACATCT 4260  
QY 4374 GATTTAAAGAAACATCTTTCAGAAAGCTGTAAAGAAACCTACCTGCGAGGAATCAAAACGAAAG 4433  
DB 4261 GATTTAAAGAAACATCTTTCAGAAAGCTGTAAAGAAACCTACCTGCGAGGAATCAAAACGAAAG 4320  
QY 4434 TTGTGTTTTGCTCTAAGTATGCTAGGAAATCTCTCAGATTACTTTGCTAGATGAACCATCT 4493  
DB 4321 TTGTGTTTTGCTCTAAGTATGCTAGGAAATCTCTCAGATTACTTTGCTAGATGAACCATCT 4380  
QY 4494 ACAGGTATGGATCCCAAGCCAAACACAGCACATGTGCGAGCAATTCGAACTGCAATTTAAA 4553  
DB 4381 ACAGGTATGGATCCCAAGCCAAACACAGCACATGTGCGAGCAATTCGAACTGCAATTTAAA 4440  
QY 4554 AACAGAAAGCGGCTGCTATTTCTGACCATCTCACTATATGAGAGGAGCAGAGCTGTCTGT 4613  
DB 4441 AACAGAAAGCGGCTGCTATTTCTGACCATCTCACTATATGAGAGGAGCAGAGCTGTCTGT 4500  
QY 4614 GATTCGATAGCTATCATGTGTCTGCGCAGTTTAAAGATGTATCGAAACAGTACAAACATCTA 4673  
DB 4501 GATTCGATAGCTATCATGTGTCTGCGCAGTTTAAAGATGTATCGAAACAGTACAAACATCTA 4560  
QY 4674 AAGAGTAAATTTGGAAAAAGCTACTTTTGGAAAAATTTAAATTTGAAGACCTGGATAGAAAAAC 4733  
DB 4561 AAGAGTAAATTTGGAAAAAGCTACTTTTGGAAAAATTTAAATTTGAAGACCTGGATAGAAAAAC 4620  
QY 4734 CTAGAAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATTCAGACCGCTCAG 4793  
DB 4621 CTAGAAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATTCAGACCGCTCAG 4680  
QY 4794 GAAAGTTTTTCTTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCA 4853

Db 4681 GAAAGTTTTTCTCTATTTTGGCTTATATAAATCTTAAGGAAGATGTTTCAGTCCCTTCA 4740  
Qy 4854 CAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCATTTGAAGATATAGCTTT 4913  
Db 4741 CAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCATTTGAAGATATAGCTTT 4800  
Qy 4914 TCTCAAGCAACATTTGGAAGAGCTTTTCTAGAACTCTCACTAAGAACAAAGAGGGAAGAT 4973  
Db 4801 TCTCAAGCAACATTTGGAAGAGCTTTTCTAGAACTCTCACTAAGAACAAAGAGGGAAGAT 4860  
Qy 4974 AATAGTTGTGGAACTTTTAAACAGCACATTTGGTGGGAACGAAACACAAAGAGATAGAT 5033  
Db 4861 AATAGTTGTGGAACTTTTAAACAGCACATTTGGTGGGAACGAAACACAAAGAGATAGAT 4920  
Qy 5034 GTATTTGA 5042  
Db 4921 GTATTTGA 4929

RESULT 6  
AAD37619  
ID AAD37619 standard; cDNA; 4785 BP.  
XX  
AC AAD37619;  
XX  
DT 10-SEP-2002 (first entry)  
XX  
DE Human transporter protein cDNA #2.  
XX  
KW Human; novel human protein; NHP; transporter protein; mental disorder;  
KW cancer; gene therapy; drug screening; nutraceutical application; gene;  
KW cosmetic application; polymorphism; ss.  
XX  
OS Homo sapiens.  
XX  
PH Key  
FT CDS  
FT Location/Qualifiers  
FT 1..4785  
FT /tag= a  
FT /product= "Human transporter protein #2"  
FT /transl\_except= (pos:808..810, aa:Tyr)  
FT /transl\_except= (pos:2494..2496, aa:Ser)  
FT /transl\_except= (pos:2878..2880, aa:Val)  
FT misc\_feature  
FT 810  
FT /tag= b  
FT /note= "This degenerate base represents a polymorphic  
FT site"  
FT 2494  
FT misc\_feature  
FT /tag= c  
FT /note= "This degenerate base represents a polymorphic  
FT site"  
FT 2878  
FT misc\_feature  
FT /tag= d  
FT /note= "This degenerate base represents a polymorphic  
FT site"  
FT 2878  
FT WO200231147-A2.  
XX  
XX 18-APR-2002.  
XX  
XX 04-OCT-2001; 2001WO-US031113.  
XX  
XX 10-OCT-2000; 2000US-0239629P.  
XX  
XX (LEXI-) LEXICON GENETICS INC.  
XX  
XX Hu Y, Nepomnichy B;  
XX  
XX WPI; 2002-454552/48.  
XX  
XX P-PSDB; AAE23657.  
XX  
XX Novel nucleic acid molecule encoding novel human proteins, useful for  
XX therapeutic, diagnostic and pharmacogenomic applications.  
PT

XX Claim 1; Page 39-41; 46pp; English.  
PS  
XX The present sequence is a cDNA coding for novel human protein (NHP),  
CC human transporter protein. NHPs shares structural similarity with the  
CC mammalian ATP-binding cassette (ABC) transporters and multidrug  
CC resistance transporters. NHP polynucleotides are useful for the  
CC therapeutic, diagnostic and pharmacogenomic applications. They are used  
CC for detecting and treating mental disorders and cancers. They are also  
CC used in gene therapy. NHP polypeptides are useful for diagnosis, drug  
CC screening, clinical trial monitoring, treatment of diseases and  
CC disorders, and cosmetic or nutraceutical applications  
XX  
SQ Sequence 4785 BP; 1489 A; 807 C; 917 G; 1568 T; 0 U; 4 Other;  
Query Match 80.5%; Score 4408; DB 6; Length 4785;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 4758; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Qy 114 ATGTCACCTGCAATTAGGGAGGTAGAGTTTGGAGACAGACAGACACTTCTACTGAG 173  
Db 1 ATGTCACCTGCAATTAGGGAGGTAGAGTTTGGAGACAGACAGACACTTCTACTGAG 60  
Qy 174 AATTACTTAATTAATGCAGAACCAAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA 233  
Db 61 AATTACTTAATTAATGCAGAACCAAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA 120  
Qy 234 TTTTCTTTTATTTGGTTAATATTAATAGCATGATGCATCCAAATGAAGAAATGAGAA 293  
Db 121 TTTTCTTTTATTTGGTTAATATTAATAGCATGATGCATCCAAATGAAGAAATGAGAA 180  
Qy 294 GTGCTTAATATAGAACTCAATCTATGGACAAAGTTTCTTCTTAATCTAATTTCTTGA 353  
Db 181 GTGCTTAATATAGAACTCAATCTATGGACAAAGTTTCTTCTTAATCTAATTTCTTGA 240  
Qy 354 TATATCTCCAGTGACTAATATTAACAGCAGCATCATGAGAAAGTGCTCTACTGATCAT 413  
Db 241 TATATCTCCAGTGACTAATATTAACAGCAGCATCATGAGAAAGTGCTCTACTGATCAT 300  
Qy 414 CCGTATGCTCATTAATTAATGAGAAATATCAAAATGAAGAAAGTATTAACATCCAGTCTC 473  
Db 301 CCGTATGCTCATTAATTAATGAGAAATATCAAAATGAAGAAAGTATTAACATCCAGTCTC 360  
Qy 474 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAGATCCATGCTCTATGAAGAACTTCT 533  
Db 361 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAGATCCATGCTCTATGAAGAACTTCT 420  
Qy 534 TTTTCTCTCATATGATCCAGTATCTTCTATTTATATGATTCAGAGCTGGCTGTCTCA 593  
Db 421 TTTTCTCTCATATGATCCAGTATCTTCTATTTATATGATTCAGAGCTGGCTGTCTCA 480  
Qy 594 AAATCATGTGAGGCTGCTCAGTACTGCTCAGTACTGCTCAGTACTGCTCAGTACTGCTC 653  
Db 481 AAATCATGTGAGGCTGCTCAGTACTGCTCAGTACTGCTCAGTACTGCTCAGTACTGCTC 540  
Qy 654 GATGCTGCCATTTATACAGTTGAAGACCAATGTTTCTTTTGGAGAGCTGGAGTCAACT 713  
Db 541 GATGCTGCCATTTATACAGTTGAAGACCAATGTTTCTTTTGGAGAGCTGGAGTCAACT 600  
Qy 714 AAAGCTGTATATGGAGAAACTGCTGTGTGTGAAGAAATAGATACCTTTTCCCGAGAGTA 773  
Db 601 AAAGCTGTATATGGAGAAACTGCTGTGTGTGAAGAAATAGATACCTTTTCCCGAGAGTA 660  
Qy 774 ATTTTAATATACCTAGTTATAGCATTTTACCTTTTGGATACCTTTTGGCAATTCATC 833  
Db 661 ATTTTAATATACCTAGTTATAGCATTTTACCTTTTGGATACCTTTTGGCAATTCATC 720  
Qy 834 GTACGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 893  
Db 721 GTACGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 780  
Qy 894 GCCTTTTGGCTTCTCCTGGTCTCTATATACAGTTTAAATTTTCTTATGCTCCTTCT 953

Db 781 GCGTTTGGCTTCTCTGAGGCTTCTCTATATAACAGTTTAAATTTTCTTATGTCCTCTT 840  
 Qy 954 ATGGCAGTCAATTCGACAGCTTCTTTGTTATTTCTTCAAGTAGCAGCATTTGTGATATT 1013  
 Db 841 ATGGCAGTCAATTCGACAGCTTCTTTGTTATTTCTTCAAGTAGCAGCATTTGTGATATT 900  
 Qy 1014 CTGCTTTTCTTCTTATGATATATCATCTGATTTATTTTGTCTTAAATGCTGACACCTCTT 1073  
 Db 901 CTGCTTTTCTTCTTATGATATATCATCTGATTTATTTTGTCTTAAATGCTGACACCTCTT 960  
 Qy 1074 TTTTAAATAATCAAAACATGCGGAAATAGTGAATTTTGTGTAATGCTGCTGCTTTGGATTT 1133  
 Db 961 TTTTAAATAATCAAAACATGCGGAAATAGTGAATTTTGTGTAATGCTGCTGCTTTGGATTT 1020  
 Qy 1134 ATTTGGCTTATGATTAATCTCATAGAAATTTTCCAAATCTGTAGTGTGCTTTTCAGT 1193  
 Db 1021 ATTTGGCTTATGATTAATCTCATAGAAATTTTCCAAATCTGTAGTGTGCTTTTCAGT 1080  
 Qy 1194 CTTTCTGTCACGTACTTCTTCTGATTTGATTTGATTTGACAGGTCATGCAATTTAGAGATTTT 1253  
 Db 1081 CTTTCTGTCACGTACTTCTTCTGATTTGATTTGATTTGACAGGTCATGCAATTTAGAGATTTT 1140  
 Qy 1254 AATGAAGTGTCTTCAATTTCAAAATTTGACTGAGGCCCATATCTCTTAATTTATACAAAT 1313  
 Db 1141 AATGAAGTGTCTTCAATTTCAAAATTTGACTGAGGCCCATATCTCTTAATTTATACAAAT 1200  
 Qy 1314 ATCATGCTCACACTTAATAGTATATTTATGTCCTCTTGGCTGTCTATCTTTGATCAAGTC 1373  
 Db 1201 ATCATGCTCACACTTAATAGTATATTTATGTCCTCTTGGCTGTCTATCTTTGATCAAGTC 1260  
 Qy 1374 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAAGCTTCATATTGG 1433  
 Db 1261 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAAGCTTCATATTGG 1320  
 Qy 1434 TCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGGCAATGTTAAATGGAATATTAGT 1493  
 Db 1321 TCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGGCAATGTTAAATGGAATATTAGT 1380  
 Qy 1494 TTTAGTGAATTTATGAGGAGTCTTCTCAGATTTGTAGGAAAGAGCCATAGAAAT 1553  
 Db 1381 TTTAGTGAATTTATGAGGAGTCTTCTCAGATTTGTAGGAAAGAGCCATAGAAAT 1440  
 Qy 1554 AGTGGTATTTCAGAGACATACAGAAAGAGGGTGAAATTTGCGAGGCTTTGAGAAATTTG 1613  
 Db 1441 AGTGGTATTTCAGAGACATACAGAAAGAGGGTGAAATTTGCGAGGCTTTGAGAAATTTG 1500  
 Qy 1614 TCATTTGACATATATGAGGGTCAGATTTACTGCTTACTTTGGCCAAGTGGAAACAGGAAAG 1673  
 Db 1501 TCATTTGACATATATGAGGGTCAGATTTACTGCTTACTTTGGCCAAGTGGAAACAGGAAAG 1560  
 Qy 1674 AGTACATTTGATGAATTTCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1733  
 Db 1561 AGTACATTTGATGAATTTCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1620  
 Qy 1734 TATGGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAAGAAATGATTTGGCAT 1793  
 Db 1621 TATGGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAAGAAATGATTTGGCAT 1680  
 Qy 1794 TGTCCACAGTTAGATATACATTTGATGTTTGAAGTAGAGAAATTTATCAATTTTG 1853  
 Db 1681 TGTCCACAGTTAGATATACATTTGATGTTTGAAGTAGAGAAATTTATCAATTTTG 1740  
 Qy 1854 GCTTCAATCAAGGGATACAGCCCAATATAATACAGAGAGTGCAGAAAGGTTTACTA 1913  
 Db 1741 GCTTCAATCAAGGGATACAGCCCAATATAATACAGAGAGTGCAGAAAGGTTTACTA 1800  
 Qy 1914 GATTTAGACATGACAGTCTCAAGATTAACCAAGCTTAAATAATTAAGTGGTGGTCAAAA 1973  
 Db 1801 GATTTAGACATGACAGTCTCAAGATTAACCAAGCTTAAATAATTAAGTGGTGGTCAAAA 1860  
 Qy 1974 AGAAGCTGTCTTAGGAATTTGCTGTTCTTTGGGAACCCAAACATATCTGCTGATGAA 2033  
 Db 1861 AGAAGCTGTCTTAGGAATTTGCTGTTCTTTGGGAACCCAAACATATCTGCTGATGAA 1920

Qy 2034 CCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAATACAGA 2093  
 Db 1921 CCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAATACAGA 1980  
 Qy 2094 AAAGCCAATCGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGATCTTTGCA 2153  
 Db 1981 AAAGCCAATCGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGATCTTTGCA 2040  
 Qy 2154 GATAGGAAGCTGTGATATACAGGAATGCTGAAATGTTGGTCTTCAATGTTCCCTC 2213  
 Db 2041 GATAGGAAGCTGTGATATACAGGAATGCTGAAATGTTGGTCTTCAATGTTCCCTC 2100  
 Qy 2214 AAAGTAAATGCGGGATCGGCTACCCCTGAGCATGTACATAGCAAAATTTTGTGCCACA 2273  
 Db 2101 AAAGTAAATGCGGGATCGGCTACCCCTGAGCATGTACATAGCAAAATTTTGTGCCACA 2160  
 Qy 2274 GAATCTCTTTCTTCACTGTTTAAACAATATACATATACCTGGAGCTACTTTATACAGAGAT 2333  
 Db 2161 GAATCTCTTTCTTCACTGTTTAAACAATATATACATATACCTGGAGCTACTTTATACAGAGAT 2220  
 Qy 2334 GACCAACAATCTGTATAGCTTGCCTTTCAAGGACATGACAAATTTTTCAGGTTTGT 2393  
 Db 2221 GACCAACAATCTGTATAGCTTGCCTTTCAAGGACATGACAAATTTTTCAGGTTTGT 2280  
 Qy 2394 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGTGTTCATGACACT 2453  
 Db 2281 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGGTTCATGACACT 2340  
 Qy 2454 TTGGAAGACGTATTTTAAAGCTAGAGTTTGAAGCAGAAATTTGACCAAGCATATTAGT 2513  
 Db 2341 TTGGAAGACGTATTTTAAAGCTAGAGTTTGAAGCAGAAATTTGACCAAGCATATTAGT 2400  
 Qy 2514 GTATTTACTCAGCAGCCTCGAGGAGAAATGGATTCAAAATCTTTTGAATGAATGGA 2573  
 Db 2401 GTATTTACTCAGCAGCCTCGAGGAGAAATGGATTCAAAATCTTTTGAATGAATGGA 2460  
 Qy 2574 CAGAGCTTACTTATCTTCTGAAACCAAGGCTTCTAGTAGGACCATGAGCCTTTGG 2633  
 Db 2461 CAGAGCTTACTTATCTTCTGAAACCAAGGCTTCTAGTAGGACCATGAGCCTTTGG 2520  
 Qy 2634 AAACAACAGATGTATCAATFAGCAAAAGTTTCATTTCTTTACCTTGAAGAGTGAAT 2693  
 Db 2521 AAACAACAGATGTATCAATFAGCAAAAGTTTCATTTCTTTACCTTGAAGAGTGAAT 2580  
 Qy 2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTT 2753  
 Db 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTT 2640  
 Qy 2754 TTGGTTCACTCACTCTTTTAAATTTGCTGTGGTTCCCATCAAACTTGTTCAGACTTAT 2813  
 Db 2641 TTGGTTCACTCACTCTTTTAAATTTGCTGTGGTTCCCATCAAACTTGTTCAGACTTAT 2700  
 Qy 2814 TTTCTAAACCTCGAGACAAACCAATATAAATAAACAAGTGTGCTTCTTCAAAATCT 2873  
 Db 2701 TTTCTAAACCTCGAGACAAACCAATATAAATAAACAAGTGTGCTTCTTCAAAATCT 2760  
 Qy 2874 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTCAAGCCAGCAACATAATGGTGA 2933  
 Db 2761 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTCAAGCCAGCAACATAATGGTGA 2820  
 Qy 2934 ATGATTAATGACAGTGTATGATCCGTGGCTCCCATAGTCGGCTTTAAATGTGATG 2993  
 Db 2821 ATGATTAATGACAGTGTATGATCCGTGGCTCCCATAGTCGGCTTTAAATGTGATG 2880  
 Qy 2994 CATTCAAAAGGACATGTTTGTGAGCTGTTTTCACAGTACTATGGTATTCTTTA 3053  
 Db 2881 CATTCAAAAGGACATGTTTGTGAGCTGTTTTCACAGTACTATGGTATTCTTTA 2940  
 Qy 3054 CCTATATTAGTGAATTCATTAGTAACTACTATCTTTATCAATTTAAATGTGACTGAACC 3113  
 Db 2941 CCTATATTAGTGAATTCATTAGTAACTACTATCTTTATCAATTTAAATGTGACTGAACC 3000

QY 3114 ATCCGATCTGGAGTACCCCATCTTTCAAGAAATCTGATATAGTTTTTAAATTCGAG 3173  
DB 3001 ATCCGATCTGGAGTACCCCATCTTTCAAGAAATCTGATATAGTTTTTAAATTCGAG 3060  
QY 3174 CTGTAATTTTCAAGCAGCTTTGCTTGGAAATCAATTTGTAATGCAATCCACCTTACTTTGCC 3233  
DB 3061 CTGTAATTTTCAAGCAGCTTTGCTTGGAAATCAATTTGTAATGCAATCCACCTTACTTTGCC 3120  
QY 3234 ATGGAAATGCGAGAAATCATAGATCAAGCTTAAAGCTTATCTCACTTAAACTTTTCAAGTCTT 3293  
DB 3121 ATGGAAATGCGAGAAATCATAGATCAAGCTTAAAGCTTAAACTTTTCAAGTCTT 3180  
QY 3294 TTGCCATCTGCAATTTGGAAATGCGAAGCTTTGTTGATATCCCTTATTTTATCATTT 3353  
DB 3181 TTGCCATCTGCAATTTGGAAATGCGAAGCTTTGTTGATATCCCTTATTTTATCATTT 3240  
QY 3354 CTATTTTGAATGCTAGGAACTTATGGCAATTTCAATTAAGATATATTTTATCTGTA 3413  
DB 3241 CTATTTTGAATGCTAGGAACTTATGGCAATTTCAATTAAGATATATTTTATCTGTA 3300  
QY 3414 AAGTTCTTGGCTGTGTTTTCCTTATTTGGTTATTTTCCATCAGTTTATTTCTGTTCACT 3473  
DB 3301 AAGTTCTTGGCTGTGTTTTCCTTATTTGGTTATTTTCCATCAGTTTATTTCTGTTCACT 3360  
QY 3474 TATATTGCTTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3533  
DB 3361 TATATTGCTTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
QY 3534 TATTTCTGTCGACGCTTTGCTTGTATTTGCAATCACTGAAATTAACCTTTCTTTATGGGATAC 3593  
DB 3421 TATTTCTGTCGACGCTTTGCTTGTATTTGCAATCACTGAAATTAACCTTTCTTTATGGGATAC 3480  
QY 3594 ACAATTTGCAATCTTCTTCAATTTGCTTTGTAATCATCATTTCCAACTTATCCACTTCTA 3653  
DB 3481 ACAATTTGCAATCTTCTTCAATTTGCTTTGTAATCATCATTTCCAACTTATCCACTTCTA 3540  
QY 3654 GGTTCCTGATTTCTTTCAATAAGATTTCTTGAAGAAATGTCGAAATAATGTCGACACC 3713  
DB 3541 GGTTCCTGATTTCTTTCAATAAGATTTCTTGAAGAAATGTCGAAATAATGTCGACACC 3600  
QY 3714 TATAATCCATGGGATAGGCTTTTCAATTTGCTTTGTAATCATCATTTCCAACTTATCCACTTCTA 3773  
DB 3601 TATAATCCATGGGATAGGCTTTTCAATTTGCTTTGTAATCATCATTTCCAACTTATCCACTTCTA 3660  
QY 3774 TGGATTTTCTCTTCAATCACTATGAGAAATATGAGGAGGAGATCAATTAAGAAATGAT 3833  
DB 3661 TGGATTTTCTCTTCAATCACTATGAGAAATATGAGGAGGAGATCAATTAAGAAATGAT 3720  
QY 3834 CCCTTTTTCAGAAACCTTTTCAACGAGTCTAAAAATAGGAAGCTTCCAGAACCCAGAC 3893  
DB 3721 CCCTTTTTCAGAAACCTTTTCAACGAGTCTAAAAATAGGAAGCTTCCAGAACCCAGAC 3780  
QY 3894 AATGAGGATGAAGATGAAGATGTCAGAACTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 3953  
DB 3781 AATGAGGATGAAGATGAAGATGTCAGAACTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 3840  
QY 3954 TGGCAGTGTGTGAGGAGAAACCATCCATTTATGTTGTCAGCAATTTGCTAATAAGAAATGAT 4013  
DB 3841 TGGCAGTGTGTGAGGAGAAACCATCCATTTATGTTGTCAGCAATTTGCTAATAAGAAATGAT 3900  
QY 4014 GACAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 4073  
DB 3901 GACAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 3960  
QY 4074 TTCTGTGTGAAAGAGAGATCTTTAGGACTTATTTGGTCCAAATGTTGCTGGCAAAAGC 4133  
DB 3961 TTCTGTGTGAAAGAGAGATCTTTAGGACTTATTTGGTCCAAATGTTGCTGGCAAAAGC 4020  
QY 4134 ACAATTTAATATTTCTGTTGTTGATTTTGAACCACTTCCAGGCGCAGGATTTTTTAGGA 4193  
DB 4021 ACAATTTAATATTTCTGTTGTTGATTTTGAACCACTTCCAGGCGCAGGATTTTTTAGGA 4080  
QY 4194 GATTATTTCTCAGACAGAGTGAAGATGATGATTTCACTGAAATGATGTTGCTGTTCTGCTCT 4253

DB 4081 GATTATTTCTCAGACAGAGTGAAGTGAATGATTTCACTGAAGTATGGTTACTGCTCT 4140  
QY 4254 CAGATAAACCTTTTGTGGCCAGATACATATTCAGAGAAATTTTGAATTTATGAGCT 4313  
DB 4141 CAGATAAACCTTTTGTGGCCAGATACATATTCAGAGAAATTTTGAATTTATGAGCT 4200  
QY 4314 GTCAAGGAATGATGTCAGAGTGCATGAAGAGTCAATGAAGTCAATGAAGTCAATGAAGTCAATGAAGT 4373  
DB 4201 GTCAAGGAATGATGTCAGAGTGCATGAAGAGTCAATGAAGTCAATGAAGTCAATGAAGTCAATGAAGT 4260  
QY 4374 GATTAAAGAAATCTTTCAGAGAACTGTAAGAAATCTACCTGCGAGAAATCAACGAAAG 4433  
DB 4261 GATTAAAGAAATCTTTCAGAGAACTGTAAGAAATCTACCTGCGAGAAATCAACGAAAG 4320  
QY 4434 TTGTGTTTGTCTCTAAGTATGCTAGGAAATCTCAGATTTACTTTGCTAGATGAACCATCT 4493  
DB 4321 TTGTGTTTGTCTCTAAGTATGCTAGGAAATCTCAGATTTACTTTGCTAGATGAACCATCT 4380  
QY 4494 ACAGGTATGGATCCCAAGCCAAACAGCACTGTGGCGAGCAATTCGAATCGATTTAA 4553  
DB 4381 ACAGGTATGGATCCCAAGCCAAACAGCACTGTGGCGAGCAATTCGAATCGATTTAA 4440  
QY 4554 AACAGAAAGCGGCTGCTATTTCTGACCACTCCTATATGAGGAGGAGGAGGCTGCTGT 4613  
DB 4441 AACAGAAAGCGGCTGCTATTTCTGACCACTCCTATATGAGGAGGAGGAGGCTGCTGT 4500  
QY 4614 GATCGAGTATCATGTTGCTGCGGAGTTAAGATGATCGGAACAGTACAACTCA 4673  
DB 4501 GATCGAGTATCATGTTGCTGCGGAGTTAAGATGATCGGAACAGTACAACTCA 4560  
QY 4674 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTTAAATTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 4733  
DB 4561 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTTAAATTTGAAGCTTGAAGCTTGAAGCTTGAAGCT 4620  
QY 4734 CTAGAAGTACAGCCCTTCAAAGAGAAATTCAGTATATTTTCCCAATTCAGAGCTGCTAG 4793  
DB 4621 CTAGAAGTACAGCCCTTCAAAGAGAAATTCAGTATATTTTCCCAATTCAGAGCTGCTAG 4680  
QY 4794 GAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGAGATGTTTCAAGCTTCTCA 4853  
DB 4681 GAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGAGATGTTTCAAGCTTCTCA 4740  
QY 4854 CAATCTTTTAAAGCTGGAAGAG 4878  
DB 4741 CAATCTTTTAAAGCTGGAAGAG 4765

RESULT 7  
ABSS7749  
ID ABSS7749 standard; cdna; 5463 BP.  
XX  
AC ABSS7749;  
XX  
DT 04-FEB-2003 (first entry)  
XX  
XX cdna encoding novel human ATP binding cassette ABCA5 transporter #1.  
Human; ATP binding cassette; ABC; ABCA5; transporter;  
neurotoxin transport; beta-amyloid peptide; chromosome mapping;  
blood brain barrier transport; tissue typing; predictive medicine;  
ABCA5 mediated disorder; ABCA5 related disorder; gene therapy; gene; ss.  
Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 114..5030  
FT CDS /\*tag= a  
FT /product= "ABCA5"  
FT /note= "ATP binding cassette (ABC) A5 transporter"  
FT /transl\_except= (pos:3273..3275, aa:Ser)  
XX  
PN US2002123107-A1.



XX 05-SEP-2002.  
 XX 01-MAR-2002; 2002US-00090458.  
 XX 02-MAR-2001; 2001US-0272885P.  
 XX (ACTI-) ACTIVE PASS PHARM INC.  
 XX Chen H, Killinski L, Le Bihan S;  
 XX WPI; 2003-066798/06.  
 XX P-PSDB; ABG72423.  
 DR Novel isolated ATP binding cassette transporter family polypeptide,  
 PT ABCA5, useful for treating disorders associated with aberrant or unwanted  
 PT ABCA5 transporter expression or activity.  
 XX Claim 2; Fig 1; 52pp; English.  
 XX The invention describes an isolated ATP binding cassette (ABC)  
 CC transporter family polypeptide (I), designated ABCA5. (I) or the  
 CC polynucleotide encoding it (II) are useful as targets for developing  
 CC modulating agents to regulate a variety of cellular processes,  
 CC particularly the transport of neurotoxic molecules, e.g., beta-amyloid  
 CC peptide (Abeta), across cell membranes or, e.g., the blood brain barrier  
 CC (BBB), as targets for developing modulating agents of multi-drug  
 CC resistance, as diagnostic and therapeutic tools, or to treat disorders  
 CC associated with aberrant or unwanted ABCA5 transporter expression or  
 CC activity. (I), (II) or a host cell (III) expressing (II) are useful in  
 CC screening assays, detection assays (e.g., chromosomal mapping, tissue  
 CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,  
 CC prognostic assays, monitoring clinical trials and pharmacogenomics), and  
 CC in methods of treatment (e.g., therapeutic and prophylactic). (I) or  
 CC (II) are useful as reagents or targets in assays applicable to treatment  
 CC and diagnosis of ABCA5-mediated or related disorders. (I) is useful to  
 CC screen for naturally occurring ABCA5 substrates; to screen for drugs or  
 CC compounds which modulate ABCA5 activity; as a bait protein in a yeast two  
 CC -hybrid or three-hybrid assay; and to identify other proteins which bind  
 CC to or interact with ABCA5. (II) is useful in: gene therapy; to detect  
 CC ABCA5 mRNA or a genetic alteration in a ABCA5 gene; to modulate ABCA5  
 CC activity; to locate gene regions associated with genetic disease or to  
 CC associate ABCA5 with the disease, to identify an individual from a minute  
 CC biological sample (tissue typing), and to aid in forensic identification  
 CC of the biological sample. This sequence encodes a novel human ATP binding  
 CC cassette (ABC) A5 transporter  
 XX  
 SQ Sequence 5463 BP; 1699 A; 912 C; 1055 G; 1794 T; 0 U; 3 Other;  
 Query Match 51.6%; Score 2826; DB 7; Length 5463;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2876; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GGGTCGGCCCTCGCAGATCCNAGCTGGTCCGCCCACTGAGTCAACAGACTCGAGC 60  
 DB 1 GGGTCGGCCCTCGCAGATCCNAGCTGGTCCGCCCACTGAGTCAACAGACTCGAGC 60  
 QY 61 GGGTCAGGCTGACAGCTCTCGGCTCGGGCCCTGAGGTTATTTCAGAAACATGTCCA 120  
 DB 61 GGGTCAGGCTGACAGCTCTCGGCTCGGGCCCTGAGGTTATTTCAGAAACATGTCCA 120  
 QY 121 CTGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGCAACACTTCTACTGAAGAATTACT 180  
 DB 121 CTGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGCAACACTTCTACTGAAGAATTACT 180  
 QY 181 TAATTAATGCAAGACCAAAAAGAGTAGTTTCAGGAAATCTTTTCCACTATTTTTT 240  
 DB 181 TAATTAATGCAAGACCAAAAAGAGTAGTTTCAGGAAATCTTTTCCACTATTTTTT 240  
 QY 241 TATTTTGGTTAATTAATAGCATGTCATCCAAATAGAAATATGAAAGAGTGCTTA 300  
 DB 241 TATTTTGGTTAATTAATAGCATGTCATCCAAATAGAAATATGAAAGAGTGCTTA 300

QY 301 ATATAGAACTCAATCCTATGGAAGAGTTTACTCTTTCTAATCTAATCTTGGATATATCTC 360  
 DB 301 ATATAGAACTCAATCCTATGGAAGAGTTTACTCTTTCTAATCTAATCTTGGATATATCTC 360  
 QY 361 CAGTGACTTAATATTACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTACCTGATG 420  
 DB 361 CAGTGACTTAATATTACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTACCTGATG 420  
 QY 421 TCATAATTACTGAAGAATATACAATAAGAAAGAAATGTTAAATCATCCAGTCTCTCTAAGC 480  
 DB 421 TCATAATTACTGAAGAATATACAATAAGAAAGAAATGTTAAATCATCCAGTCTCTCTAAGC 480  
 QY 481 CGAGCAACTTTGTAGGTGTGTTTCAAAGACTCCCATGTCTTATGAACCTTCGTTTTTTC 540  
 DB 481 CGAGCAACTTTGTAGGTGTGTTTCAAAGACTCCCATGTCTTATGAACCTTCGTTTTTTC 540  
 QY 541 CTGATATGATTCGAGTATCTTCTATTATATATGATTTCAAGAGCTGCTCTTCAAAATCAT 600  
 DB 541 CTGATATGATTCGAGTATCTTCTATTATATATGATTTCAAGAGCTGCTCTTCAAAATCAT 600  
 QY 601 GTGAGGCTGCTCAGTACTGCTCTCAGGTTTCAAGCTTTTACAAGCATCCATAGATGCTG 660  
 DB 601 GTGAGGCTGCTCAGTACTGCTCTCAGGTTTCAAGCTTTTACAAGCATCCATAGATGCTG 660  
 QY 661 CCATTATACAGTTGAAAGCAATGTTTCTTTGGAAGAGCTGGAGTCAACTAAAGCTG 720  
 DB 661 CCATTATACAGTTGAAAGCAATGTTTCTTTGGAAGAGCTGGAGTCAACTAAAGCTG 720  
 QY 721 TTATTATGGGAGAACTGCTGTTGTAGAATAGATACCTTCCCGAGAGAGTAATTTAA 780  
 DB 721 TTATTATGGGAGAACTGCTGTTGTAGAATAGATACCTTCCCGAGAGAGTAATTTAA 780  
 QY 781 TATACCTAGTTATPAGCATTTTCACTTTTGGATPACTTTTGGCAATTCATATCGTAGCAG 840  
 DB 781 TATACCTAGTTATPAGCATTTTCACTTTTGGATPACTTTTGGCAATTCATATCGTAGCAG 840  
 QY 841 AAAAGAAAAAATAAAGAAATTTTAAAGTAATGGAGCTTCAATGATCTGCTTTT 900  
 DB 841 AAAAGAAAAAATAAAGAAATTTTAAAGTAATGGAGCTTCAATGATCTGCTTTT 900  
 QY 901 GGGTTTCTGGGTCTTCTATATACAAAGTTTAAATTTTCTATGTCCCTTCTATGGCAG 960  
 DB 901 GGGTTTCTGGGTCTTCTATATACAAAGTTTAAATTTTCTATGTCCCTTCTATGGCAG 960  
 QY 961 TCATTGGCAGAGCTCTTTGTTTATTTCTTCAAGTAGCAGCATTTGATATCTGCTTT 1020  
 DB 961 TCATTGGCAGAGCTCTTTGTTTATTTCTTCAAGTAGCAGCATTTGATATCTGCTTT 1020  
 QY 1021 TTTTCTTTTATGATTTATCATCTGATTTTCTTTTAAATGCTGACACCTCTTTTAAAA 1080  
 DB 1021 TTTTCTTTTATGATTTATCATCTGATTTTCTTTTAAATGCTGACACCTCTTTTAAAA 1080  
 QY 1081 AATCAAAAACATGTGGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTTATTTGGCC 1140  
 DB 1081 AATCAAAAACATGTGGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTTATTTGGCC 1140  
 QY 1141 TTATGATATCTCATAGAAAGTTTCCCAATCGTTAGTGGCTTTTCAAGTCTTTCT 1200  
 DB 1141 TTATGATATCTCATAGAAAGTTTCCCAATCGTTAGTGGCTTTTCAAGTCTTTCT 1200  
 QY 1201 GTCACTGTACTTTTGTGATTGTTATGCAAGCTCATGCAATTGAGAATTTTAAATGAAG 1260  
 DB 1201 GTCACTGTACTTTTGTGATTGTTATGCAAGCTCATGCAATTGAGAATTTTAAATGAAG 1260  
 QY 1261 GTGCTTCATTTTCAAAATTTGACGAGGCCCATATCTCTTAATTTATCAATATCATGTC 1320  
 DB 1261 GTGCTTCATTTTCAAAATTTGACGAGGCCCATATCTCTTAATTTATCAATATCATGTC 1320  
 QY 1321 TCACACTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTGAATCAAGTCAATCCAG 1380  
 DB 1321 TCACACTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTGAATCAAGTCAATCCAG 1380  
 QY 1381 GGGAAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCCTTCATATTTGTCAGAGA 1440



Db 1381 GGAATTTGGCTTACGGAGATCATCTTTATATTTCTGAAGCCTTCATATTGGTCAAGA 1440  
QY 1441 GCAAAAGAAATATGAGGAGTTTATCAGAGGGCAATGTTAATGGAATATTAGTTTACTG 1500  
Db 1441 GCAAAAGAAATATGAGGAGTTTATCAGAGGGCAATGTTAATGGAATATTAGTTTACTG 1500  
QY 1501 AATATTATGAGCAGTTCTTCAGAAATTTGTAGGAAGAGCCATAGAATTTAGTGTA 1560  
Db 1501 AATATTATGAGCAGTTCTTCAGAAATTTGTAGGAAGAGCCATAGAATTTAGTGTA 1560  
QY 1561 TTCAAGAAGACATACAGAAAGAGGGTGAAATTTGTGAGGCTTTGAGAAATTTGCAATTTG 1620  
Db 1561 TTCAAGAAGACATACAGAAAGAGGGTGAAATTTGTGAGGCTTTGAGAAATTTGCAATTTG 1620  
QY 1621 ACATATATGAGGGTCAGATTTCTGCTTACCTTTGGCCACAGTGGACAGGAAGAGTACAT 1680  
Db 1621 ACATATATGAGGGTCAGATTTCTGCTTACCTTTGGCCACAGTGGACAGGAAGAGTACAT 1680  
QY 1681 TGATGAATATTTCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCACTATATATGGAC 1740  
Db 1681 TGATGAATATTTCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCACTATATATGGAC 1740  
QY 1741 ACAGAGTCTCAGAAATAGATGAATTTGTAAGCAAGAAATGATGGCAATTTGTCCAC 1800  
Db 1741 ACAGAGTCTCAGAAATAGATGAATTTGTAAGCAAGAAATGATGGCAATTTGTCCAC 1800  
QY 1801 AGTTAGATATACATTTGATGTTTTCAGTAGAAGAAATTTATCAATTTTGGCTTCAA 1860  
Db 1801 AGTTAGATATACATTTGATGTTTTCAGTAGAAGAAATTTATCAATTTTGGCTTCAA 1860  
QY 1861 TCAAGGGATACAGCCCAATATAATAACAAGAGTGCAGAGGTTTATCTAGATTTAG 1920  
Db 1861 TCAAGGGATACAGCCCAATATAATAACAAGAGTGCAGAGGTTTATCTAGATTTAG 1920  
QY 1921 ACATGAGACTATCAAGATACCAAGCTTAAATAATTAAGTGGTCTCAAAAAGAGAGC 1980  
Db 1921 ACATGAGACTATCAAGATACCAAGCTTAAATAATTAAGTGGTCTCAAAAAGAGAGC 1980  
QY 1981 TGTCATTTAGGAATGCTGTTCTTGGGAACCCAAAGATACGTGCTAGATGACCAACAG 2040  
Db 1981 TGTCATTTAGGAATGCTGTTCTTGGGAACCCAAAGATACGTGCTAGATGACCAACAG 2040  
QY 2041 CTGGAATGAGCCCTGTTCTGCACATATTTGATGAAATCTTTTAAATACAGAAAGCCA 2100  
Db 2041 CTGGAATGAGCCCTGTTCTGCACATATTTGATGAAATCTTTTAAATACAGAAAGCCA 2100  
QY 2101 ATCGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTTGCAATAGGA 2160  
Db 2101 ATCGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTTGCAATAGGA 2160  
QY 2161 AGCTGTGATATCAAGGAATGCTGAATGTTGGTTCTTCAATGTTCTCCTCAAAAGTA 2220  
Db 2161 AGCTGTGATATCAAGGAATGCTGAATGTTGGTTCTTCAATGTTCTCCTCAAAAGTA 2220  
QY 2221 AATGGGGATCGCTACCCCTGAGCATGTACATAGACAAATTTGTCACAGATCTC 2280  
Db 2221 AATGGGGATCGCTACCCCTGAGCATGTACATAGACAAATTTGTCACAGATCTC 2280  
QY 2281 TTTCTTCACTGTTAAACAACATATACCTGGAGCTACTTTATTAACAAGATGACCAAC 2340  
Db 2281 TTTCTTCACTGTTAAACAACATATACCTGGAGCTACTTTATTAACAAGATGACCAAC 2340  
QY 2341 ACCTTGATAGCTTGCCCTTTCAAGGACATGCAAAATTTTCAAGTTTGTCTGCCC 2400  
Db 2341 ACCTTGATAGCTTGCCCTTTCAAGGACATGCAAAATTTTCAAGTTTGTCTGCCC 2400  
QY 2401 TAGACAGTCATTCAAATTTGGGTGTCATTTCTATGTTGTTTCCATGACGATTTGGAAG 2460  
Db 2401 TAGACAGTCATTCAAATTTGGGTGTCATTTCTATGTTGTTTCCATGACGATTTGGAAG 2460  
QY 2461 ACCTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCATATAGTGTATTTA 2520

Db 2461 ACCTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCATATAGTGTATTTA 2520  
QY 2521 CTGAGCAGCCACTGGAGGAAGAAATCGATTCAAAATCTTTTGTGAAATGGAACAGAGCT 2580  
Db 2521 CTGAGCAGCCACTGGAGGAAGAAATCGATTCAAAATCTTTTGTGAAATGGAACAGAGCT 2580  
QY 2581 TACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTTGAAACAAC 2640  
Db 2581 TACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTTGAAACAAC 2640  
QY 2641 AGATGATACAAATAGCAAAAGTTTCATTTTCTTACCTTGAAGAGTAAATCAAGTGA 2700  
Db 2641 AGATGATACAAATAGCAAAAGTTTCATTTTCTTACCTTGAAGAGTAAATCAAGTGA 2700  
QY 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTTTGGTTC 2760  
Db 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTTTGGTTC 2760  
QY 2761 ATCAGTCTTTTAAAAATGCTGGTTCCCATCAAACTTGTCCAGACTTATATTTCTAA 2820  
Db 2761 ATCAGTCTTTTAAAAATGCTGGTTCCCATCAAACTTGTCCAGACTTATATTTCTAA 2820  
QY 2821 AACCTGGAGACAAACCAATATAACAAGTCTGCTTTCTTCAAAATTTCTGCTG 2877  
Db 2821 AACCTGGAGACAAACCAATATAACAAGTCTGCTTTCTTCAAAATTTCTGCTG 2877

RESULT 8  
ADA53416  
ID ADA53416 standard; cDNA; 3268 BP.  
XX  
AC ADA53416;  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Human coding sequence, SEQ ID 984.  
XX  
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN EPI293569-A2.  
XX  
PD 19-MAR-2003.  
XX  
PF 21-MAR-2002; 2002EP-00006586.  
XX  
PR 14-SEP-2001; 2001JP-00328381.  
XX  
PR 24-JAN-2002; 2002US-0350435P.  
XX  
PA (HELI-) HELIX RES INST.  
PA (SEAS-) RES ASSOC BIOTECHNOLOGY.  
XX  
PI Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
XX  
DR WPI; 2003-395539/38.  
DR P-FSD; ADA55055.  
XX  
PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
PT and/or membrane proteins, useful for developing medicines for diseases in  
PT which the gene is involved, or as target molecules for gene therapy.  
XX  
PS Claim 1; SEQ ID NO 984; 205pp; English.  
XX  
CC The present invention relates to novel human secretory or membrane  
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-  
CC ADA54071). The coding sequences are useful in the gene therapy of  
CC diseases caused by abnormalities of the proteins, e.g. cancer,  
CC inflammatory diseases, osteoporosis or neurological disease.

XX	Sequence	3268 BP; 1022 A; 540 C; 622 G; 1084 T; 0 U; 0 Other;	
SQ	Query Match	50.0%; Score 2739; DB 7; Length 3268;	
	Best Local Similarity	99.8%; Pred. No. 0;	
	Matches 3039; Conservative	0; Mismatches 6; Indels 0; Gaps 0;	
QY	2386	GTTCCTTTCTCCCTAGACAGCTCATCAAAATTTGGGTGTGCTATTTCTTTATGGTGTGTTCCA	2445
DB	220	GTTCCTTTCTCCCTAGACAGCTCATCAAAATTTGGGTGTGCTATTTCTTTATGGTGTGTTCCA	279
QY	2446	TGACGACTTTGGAAGAGCTGATTTTAAAGCTAGAAAGTTGAAGCAGAAATGACCAAGCAG	2505
DB	280	TGACGACTTTGGAAGAGCTGATTTTAAAGCTAGAAAGTTGAAGCAGAAATGACCAAGCAG	339
QY	2506	ATTATAGTGTATTTACTCAGCAGCAGCTGAGGAGAAATGAAATCAAAATCTTTTGATG	2565
DB	340	ATTATAGTGTATTTACTCAGCAGCAGCTGAGGAGAAATGAAATCAAAATCTTTTGATG	399
QY	2566	AAATGGAACAGAGCTTACTTATTTCTTGAACCAAGGCTTCTAGTGAGCACCATGA	2625
DB	400	AAATGGAACAGAGCTTACTTATTTCTTGAACCAAGGCTTCTAGTGAGCACCATGA	459
QY	2626	GCCTTTGGAACCAACAGATGATCAATAGCAAGTTTCATTTCTTTACCTTGAAACGTG	2685
DB	460	GCCTTTGGAACCAACAGATGATCAATAGCAAGTTTCATTTCTTTACCTTGAAACGTG	519
QY	2686	AAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTT	2745
DB	520	AAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTT	579
QY	2746	TTATGTTTTGTTCATCACTCTTTTAAATGCTGTGTTCCCATCAAACTTGTTCAG	2805
DB	580	TTATGTTTTGTTCATCACTCTTTTAAATGCTGTGTTCCCATCAAACTTGTTCAG	639
QY	2806	ACTTATATTTCTTAAACCTGGAGCAACACACATATATACAAACAGTCTGCTTCTC	2865
DB	640	ACTTATATTTCTTAAACCTGGAGCAACACACATATATACAAACAGTCTGCTTCTC	699
QY	2866	AAAAATCTGCTGACTCAGATATCAGTGTATCTTATTAGCTTTTTCACAAAGCAGAACATAA	2925
DB	700	AAAAATCTGCTGACTCAGATATCAGTGTATCTTATTAGCTTTTTCACAAAGCAGAACATAA	759
QY	2926	TGGTGACGATGATTAATGACAGTACTATGATCCGTGGCTCCCATAGTGGCTTTAA	2985
DB	760	TGGTGACGATGATTAATGACAGTACTATGATCCGTGGCTCCCATAGTGGCTTTAA	819
QY	2986	ATGTGATGCAATCAGAAAGGACTATGTTTTTGACAGCTGTTTTCAACAGTACTATGTTTT	3045
DB	820	ATGTGATGCAATCAGAAAGGACTATGTTTTTGACAGCTGTTTTCAACAGTACTATGTTTT	879
QY	3046	ATTCTTTACCTATATAGTGAATATCATAGTACTATCTTTTATCATTTAAATGTA	3105
DB	880	ATTCTTTACCTATATAGTGAATATCATAGTACTATCTTTTATCATTTAAATGTA	939
QY	3106	CTGAAACCATCCAGATCTGGAGTACCCATCTTTCAAGAAATTAATCATATAGTTTTTA	3165
DB	940	CTGAAACCATCCAGATCTGGAGTACCCATCTTTCAAGAAATTAATCATATAGTTTTTA	999
QY	3166	AAATTGAGCTGATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACT	3225
DB	1000	AAATTGAGCTGATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACT	1059
QY	3226	ACTTTGGCAATGGAATGCAAGGATCATAGATCAAGGCTTATCTCAACTTAACTTT	3285
DB	1060	ACTTTGGCAATGGAATGCAAGGATCATAGATCAAGGCTTATCTCAACTTAACTTT	1119
QY	3286	CAGGTCTTTTGGCATCTGCATATTTGGATTTGGCAAGCTGTTGTTGATATCCCTTATTTT	3345
DB	1120	CAGGTCTTTTGGCATCTGCATATTTGGATTTGGCAAGCTGTTGTTGATATCCCTTATTTT	1179
QY	3346	TTATCATCTTTATTTTGCTAGCAAGCTTATGGCAATTTTCATATGGATATATTTT	3405

DB	1180	TTATCATCTTATTTTGTATGCTAGGAAGCTTACTGGCAATTTCTATTATGGATTATATTTT	1239
QY	3406	ATACTGTAAAGTTCCCTGCTGTGGTATTTTGGCTTATTTGGTATTTGTCATCAGTTATTC	3465
DB	1240	ATACTGTAAAGTTCCCTGCTGTGGTATTTTGGCTTATTTGGTATTTGTCATCAGTTATTC	1299
QY	3466	TGTTCACTTATATTTGCTTCTTTCCACCTTTAAGAAAAATTTTAAATACCAAGAAATTTTGGT	3525
DB	1300	TGTTCACTTATATTTGCTTCTTTCCACCTTTAAGAAAAATTTTAAATACCAAGAAATTTTGGT	1359
QY	3526	CATTATCTATTTCTGTGGCAGCGTTGGCTTGATNTGCAATCACTGAAATAAATCTTTCTTA	3585
DB	1360	CATTATCTATTTCTGTGGCAGCGTTGGCTTGATNTGCAATCACTGAAATAAATCTTTCTTA	1419
QY	3586	TGGGATACCAATTTGCAACTATTTCTTCAATTTGTCCTTTTGTATCATCTTCAATCTATC	3645
DB	1420	TGGGATACCAATTTGCAACTATTTCTTCAATTTGTCCTTTTGTATCATCTTCAATCTATC	1479
QY	3646	CACCTCTAGTTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGTCGAAAAAATG	3705
DB	1480	CACCTCTAGTTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGTCGAAAAAATG	1539
QY	3706	TGGACACCTTATATCCATGGATAGGCTTTTCACTAGCTGTTATATCGCCTTACCTCAGT	3765
DB	1540	TGGACACCTTATATCCATGGATAGGCTTTTCACTAGCTGTTATATCGCCTTACCTCAGT	1599
QY	3766	GTCTACTGTGGATTTTCTCTTCAATACTATGAGAAAAAATATGAGGCGAGATCAATAA	3825
DB	1600	GTCTACTGTGGATTTTCTCTTCAATACTATGAGAAAAAATATGAGGCGAGATCAATAA	1659
QY	3826	GAAAAATGCCCTTTTTCAGAAACCTTTTCAACGAAGTCTAAAAATAGGAGCTTCCAGAC	3885
DB	1660	GAAAAATGCCCTTTTTCAGAAACCTTTTCAACGAAGTCTAAAAATAGGAGCTTCCAGAC	1719
QY	3886	CACGACCAATGAGGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT	3945
DB	1720	CACGACCAATGAGGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT	1779
QY	3946	TGATGGTTGCCAGTGTGTGGAGGAAACCAATCCATTTAGTTCAGCAATTTGTCATAAG	4005
DB	1780	TGATGGTTGCCAGTGTGTGGAGGAAACCAATCCATTTAGTTCAGCAATTTGTCATAAG	1839
QY	4006	AATATGATGACAAAGAAATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTAAAGTAAAGT	4065
DB	1840	AATATGATGACAAAGAAATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTAAAGTAAAGT	1899
QY	4066	ACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCAAATTTGGTGTG	4125
DB	1900	ACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCAAATTTGGTGTG	1959
QY	4126	GCAAGACCAATTTATTAATTTCTGGTTGGTGATATTGAACCACTTCAGGCCAGGTAT	4185
DB	1960	GCAAGACCAATTTATTAATTTCTGGTTGGTGATATTGAACCACTTCAGGCCAGGTAT	2019
QY	4186	TTTTAGGAGATTTATTTCTCAGAGACAAGTGAAGATGATGATTCATGAAAGTGTATGGT	4245
DB	2020	TTTTAGGAGATTTATTTCTCAGAGACAAGTGAAGATGATGATTCATGAAAGTGTATGGT	2079
QY	4246	ACTGTCTCAGATAAACCTTTTGGCCAGATCTACTATTCAGGAGCAATTTTGAATTT	4305
DB	2080	ACTGTCTCAGATAAACCTTTTGGCCAGATCTACTATTCAGGAGCAATTTTGAATTT	2139
QY	4306	ATGGAGCTGTCAAGGAATGAGTGAAGTGAACATGAAAGAAAGTCAATGCTGAAATACAC	4365
DB	2140	ATGGAGCTGTCAAGGAATGAGTGAAGTGAACATGAAAGAAAGTCAATGCTGAAATACAC	2199
QY	4366	ATGCATTTGATTTTAAAGAAACATCTTCAGAGACTGTAAAGAACTACCTGCGGATCA	4425
DB	2200	ATGCATTTGATTTTAAAGAAACATCTTCAGAGACTGTAAAGAACTACCTGCGGATCA	2259
QY	4426	AACGAAGTTGTGTTTTGCTTAAGTATGCTAGGGAATCCTCAGATTTACTTTGCTAGATG	4485
DB	2260	AACGAAGTTGTGTTTTGCTTAAGTATGCTAGGGAATCCTCAGATTTACTTTGCTAGATG	2319





QY 2334 GACCAACACTTGTGTATAGCTTGGCTTCAAGGACATGGACAAATTTTCAGGTTTGT 2393  
DB 2221 GACCAACACTTGTGTATAGCTTGGCTTCAAGGACATGGACAAATTTTCAGGTTTGT 2280  
QY 2394 TCTGCCCTAGACAGTCATTCAAAATTTGGGTGTCAATTTCTATGGTGTCTTCCATGACGACT 2453  
DB 2281 TCTGCCCTAGACAGTCATTCAAAATTTGGGTGTCAATTTCTATGGTGTCTTCCATGACGACT 2340  
QY 2454 TTGGAAGAGCTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATAGT 2513  
DB 2341 TTGGAAGAGCTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATAGT 2400  
QY 2514 GTATTACTCAGCAGCCTGAGGAGAAATGGATTCAAAATCTTTTGTGTAATGCAA 2573  
DB 2401 GTATTACTCAGCAGCCTGAGGAGAAATGGATTCAAAATCTTTTGTGTAATGCAA 2460  
QY 2574 CAGAGCTTACTTATTTCTGAAACCAAGGCTTCTAGTGAGCACCATGAGCCTTTGG 2633  
DB 2461 CAGAGCTTACTTATTTCTGAAACCAAGGCTTCTAGTGAGCACCATGAGCCTTTGG 2520  
QY 2634 AAACAACAGATGATACATAGCAAAAGTTTCATTTCTTTTACCTGAAACGTAAGTAA 2693  
DB 2521 AAACAACAGATGATACATAGCAAAAGTTTCATTTCTTTTACCTGAAACGTAAGTAA 2580  
QY 2694 TCAGTGAGATCAGTGTCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTT 2753  
DB 2581 TCAGTGAGATCAGTGTCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTT 2640  
QY 2754 TTGGTTCATCCTTTTAAATCTGTGGTTCCTCATCACTTTCCTGAGCCTTATAT 2813  
DB 2641 TTGGTTCATCCTTTTAAATCTGTGGTTCCTCATCACTTTCCTGAGCCTTATAT 2700  
QY 2814 TTCTTAAACCTGGAGACAAACCATATAAATACAAACAAAGTCTGCTTCTTCAAAATCT 2873  
DB 2701 TTCTTAAACCTGGAGACAAACCATATAAATACAAACAAAGTCTGCTTCTTCAAAATCT 2760  
QY 2874 GCTG 2877  
DB 2761 GCTG 2764

RESULT 10  
ADC51606  
ID ADC51606 standard; DNA; 2723 BP.  
XX  
AC ADC51606;  
XX  
DT 18-DEC-2003 (first entry)  
XX  
DE Human macroprotein-94.64 coding sequence, SEQ ID 1.  
XX  
KW Human; macroprotein-94.64; dementia; arrhythmia; asthma; diabetes; gene;  
XX GB.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
XX CDS 12..2485  
XX /\*tag= a  
XX /\*product= "Human macroprotein-94.64"  
XX  
XX CN1382725-A.  
XX  
XX 04-DEC-2002.  
XX  
XX 26-APR-2001; 2001CN-00112765.  
XX  
XX 26-APR-2001; 2001CN-00112765.  
XX (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.  
XX  
XX Mao Y, Xie Y;  
XX

DR WPI; 2003-269487/27.  
XX P-PSDB; ADC51607.  
PT New human macroprotein-94.64, encoding polynucleotide, antagonist and  
PT recombinant production, useful for treating dementia, arrhythmia, asthma  
PT and diabetes.  
XX  
PS Claim 6; SEQ ID NO 1; 35pp; Chinese.  
XX  
CC The present invention relates to human macroprotein-94.64 (ADC51607) and  
CC its coding sequence (ADC51606). The protein is useful for treating  
CC dementia, arrhythmia, asthma and diabetes.  
XX  
SQ Sequence 2723 BP; 855 A; 460 C; 518 G; 890 T; 0 U; 0 Other;  
  
Query Match 48.1%; Score 2633; DB 9; Length 2723;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2683; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2557 CTTTTCATGAAATGGAAACAGAGCTTACTTATTCTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
DB 1 CTTTTCATGAAATGGAAACAGAGCTTACTTATTCTTCTGAAACCAAGGCTTCTCTAGTGA 60  
QY 2617 GCACCATGAGCCTTTGGAACCAACAGATGATACATAGCAAAAGTTTCATTTCTTTACCT 2676  
DB 61 GCACCATGAGCCTTTGGAACCAACAGATGATACATAGCAAAAGTTTCATTTCTTTACCT 120  
QY 2677 TGAACGTCGAAAGTAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736  
DB 121 TGAACGTCGAAAGTAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 180  
QY 2737 TCAGATTTTATGTTTGGTTTCATCCTCTTTTAAATGCTGTGGTTCCTCATCAAC 2796  
DB 181 TCAGATTTTATGTTTGGTTTCATCCTCTTTTAAATGCTGTGGTTCCTCATCAAC 240  
QY 2797 TTGTTCCAGACTTATATTTCTAAACCTGGAGACAAACCATATAAATACAAACAAAGTC 2856  
DB 241 TTGTTCCAGACTTATATTTCTAAACCTGGAGACAAACCATATAAATACAAACAAAGTC 300  
QY 2857 TGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 2916  
DB 301 TGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 360  
QY 2917 AGAACAATATGCTGAGATGATTATTAATGACAGTACTATGTCGCTGGCTCCCATAGTG 2976  
DB 361 AGAACAATATGCTGAGATGATTATTAATGACAGTACTATGTCGCTGGCTCCCATAGTG 420  
QY 2977 CGGCTTTAAATGTGATGCAATTCAGAAAAGCACTATGTTTTCAGAGCTGTTTTCACAGTA 3036  
DB 421 CGGCTTTAAATGTGATGCAATTCAGAAAAGCACTATGTTTTCAGAGCTGTTTTCACAGTA 480  
QY 3037 CTATGGTTTATCTTTTACCTATATTAGTGAATATCATTAGTACTACTACTTATCATTT 3096  
DB 481 CTATGGTTTATCTTTTACCTATATTAGTGAATATCATTAGTACTACTACTTATCATTT 540  
QY 3097 TAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATTTCTTTTCAAGAAATTTACTGATA 3156  
DB 541 TAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATTTCTTTTCAAGAAATTTACTGATA 600  
QY 3157 TAGTTTTTAAATGTGCTGATTTTTCAGAGCCTTTGCTTGAATCATTTGTTTACTGCAA 3216  
DB 601 TAGTTTTTAAATGTGCTGATTTTTCAGAGCCTTTGCTTGAATCATTTGTTTACTGCAA 660  
QY 3217 TGCACCTTACTTTTGCATGGAATAATGCAGAGATCATTAAGTCAAAAGCTTATATCTCAAC 3276  
DB 661 TGCACCTTACTTTTGCATGGAATAATGCAGAGATCATTAAGTCAAAAGCTTATATCTCAAC 720  
QY 3277 TTAAACTTTCAGGCTTTTGCATCTGCAATTTGGATTTGGACAGCTGTTGTTGATATCC 3336  
DB 721 TTAAACTTTCAGGCTTTTGCATCTGCAATTTGGATTTGGACAGCTGTTGTTGATATCC 780  
QY 3337 CCTTATTTTATCATCTTTTATTTTGTATGCTAGGAGCTTATTTGGCATTTTCATTATGGAT 3396

Db 781 CCTATATTTTATCATCTTATTTTGATGCTAGGAAGCTTACTGGCAATTCATATGAT 840  
Qy 3397 TATATTTTATATCTGTAAAGTTCCTGTCTGTGGTTTTTGCTTATTTGGTTATGTTCCAT 3456  
Db 841 TATATTTTATATCTGTAAAGTTCCTGTCTGTGGTTTTTGCTTATTTGGTTATGTTCCAT 900  
Qy 3457 CAGTTTATCTGTTCATCTATATGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAG 3516  
Db 901 CAGTTTATCTGTTCATCTATATGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAG 960  
Qy 3517 AATTTTGGTCATTTATCTATTTCTGTGGCAGGTTGGCTTGTATTGCCAATCACTGAAATAA 3576  
Db 961 AATTTTGGTCATTTATCTATTTCTGTGGCAGGTTGGCTTGTATTGCCAATCACTGAAATAA 1020  
Qy 3577 CTTTCTTTATGGGATACAAATGGCAATNTCTTCAATATGCTTTTGTATGCAATTC 3636  
Db 1021 CTTTCTTTATGGGATACAAATGGCAATNTCTTCAATATGCTTTTGTATGCAATTC 1080  
Qy 3637 CAATCTATCCACTTCTAGGTTGGCTGATTTCTTTCATAAAGATTTCTTGGGAAGATGTAC 3696  
Db 1081 CAATCTATCCACTTCTAGGTTGGCTGATTTCTTTCATAAAGATTTCTTGGGAAGATGTAC 1140  
Qy 3697 GAAATAATGTGACACCTTATATCCATGGGATAGCTTTTCAGTAGCTGTTATATCGCCTT 3756  
Db 1141 GAAATAATGTGACACCTTATATCCATGGGATAGCTTTTCAGTAGCTGTTATATCGCCTT 1200  
Qy 3757 ACCTGCAGTGTACTGTGGATTTTCTTCTTACAAATATGAGAAAAATATGGAGCA 3816  
Db 1201 ACCTGCAGTGTACTGTGGATTTTCTTCTTACAAATATGAGAAAAATATGGAGCA 1260  
Qy 3817 GATCAATAGAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGCTTAAATAAGGAAGC 3876  
Db 1261 GATCAATAGAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGCTTAAATAAGGAAGC 1320  
Qy 3877 TTCCAGACCCACGACATGAGATGAGATGAGATGTCAGAGCTTGAAGCTTAAGG 3936  
Db 1321 TTCCAGACCCACGACATGAGATGAGATGAGATGTCAGAGCTTGAAGCTTAAGG 1380  
Qy 3937 TCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCATTCATATGTCAGCAATT 3996  
Db 1381 TCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAACCATTCATATGTCAGCAATT 1440  
Qy 3997 TGCTAAGAATATGATGACAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAGTGG 4056  
Db 1441 TGCTAAGAATATGATGACAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAGTGG 1500  
Qy 4057 CAATTAATACATCTTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGTCCAA 4116  
Db 1501 CAATTAATACATCTTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGTCCAA 1560  
Qy 4117 ATGTCGTGGCAAGCAATTAATATCTTGGTTGGTGATTAATGAAACCAACTTCAG 4176  
Db 1561 ATGTCGTGGCAAGCAATTAATATCTTGGTTGGTGATTAATGAAACCAACTTCAG 1620  
Qy 4177 GCCAGTATTTTATAGGAGATTTCTTCAGAGACAAAGTGAAGATGATTCATCTGAAGT 4236  
Db 1621 GCCAGTATTTTATAGGAGATTTCTTCAGAGACAAAGTGAAGATGATTCATCTGAAGT 1680  
Qy 4237 GTATGGGTTACTGTCCTCAGATAAACCTTTGTGGCCAGATACTACATTCAGGAACATT 4296  
Db 1681 GTATGGGTTACTGTCCTCAGATAAACCTTTGTGGCCAGATACTACATTCAGGAACATT 1740  
Qy 4297 TTGAATTTATGGAGCTGTCAGGAATGATGTCGACATGCAAGAGATCATAGTC 4356  
Db 1741 TTGAATTTATGGAGCTGTCAGGAATGATGTCGACATGCAAGAGATCATAGTC 1800  
Qy 4357 GAATAACACATGCACTTGATTTAAAGAACATCTTCAGAGAGCTGTAAAGAACTACCTG 4416  
Db 1801 GAATAACACATGCACTTGATTTAAAGAACATCTTCAGAGAGCTGTAAAGAACTACCTG 1860  
Qy 4417 CAGGAATCAACGAAGTTGTTTCTCTCTAGTATGCTAGGGAATCTCAGATTAATT 4476  
Db 1861 CAGGAATCAACGAAGTTGTTTCTCTCTAGTATGCTAGGGAATCTCAGATTAATT 1920

Qy 4477 TGCTAGATGAACCATCTACAGGTATGATCCCAAGCCCAACAGACACATGTGGCGAGCAA 4536  
Db 1921 TGCTAGATGAACCATCTACAGGTATGATCCCAAGCCCAACAGACACATGTGGCGAGCAA 1980  
Qy 4537 TTGGAATCTGATTTAAAAACAGAAAGCGGCTGCTATTCTGACCACTCATATATGAGG 4596  
Db 1981 TTGGAATCTGATTTAAAAACAGAAAGCGGCTGCTATTCTGACCACTCATATATGAGG 2040  
Qy 4597 AGCAGAGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGGCAAGTAAAGATGATCG 4656  
Db 2041 AGCAGAGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGGCAAGTAAAGATGATCG 2100  
Qy 4657 GAACTGACAAATCTAAAGAGTAAATTTGGAAAAAGGCTACTTTTTTGGAAAAATTAATGA 4716  
Db 2101 GAACTGACAAATCTAAAGAGTAAATTTGGAAAAAGGCTACTTTTTTGGAAAAATTAATGA 2160  
Qy 4717 AGGACTGGATAGAAACCTAGAAAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTCC 4776  
Db 2161 AGGACTGGATAGAAACCTAGAAAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTCC 2220  
Qy 4777 CAAATGCAAGCGCTCAGGAAAGTTTTTCTTATTTTGGCTTATAAAATTCCTAAGGAAG 4836  
Db 2221 CAAATGCAAGCGCTCAGGAAAGTTTTTCTTATTTTGGCTTATAAAATTCCTAAGGAAG 2280  
Qy 4837 ATGTTCAAGTCCCTTTCACAAATCTTTTTTAAGCTGGAAGAGCTAAACATGCTTTTGCCA 4896  
Db 2281 ATGTTCAAGTCCCTTTCACAAATCTTTTTTAAGCTGGAAGAGCTAAACATGCTTTTGCCA 2340  
Qy 4897 TTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAG 4956  
Db 2341 TTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAG 2400  
Qy 4957 AACAGAGGAGGAAGATATAGTCTGGAACCTTTAAACAGCACACTTTGGTGGGAACGA 5016  
Db 2401 AACAGAGGAGGAAGATATAGTCTGGAACCTTTAAACAGCACACTTTGGTGGGAACGA 2460  
Qy 5017 CACAAGAGATAGAGTAGTATTTTGAATTTGATTTGTTTGGTCTGCTTACTGGGACTTCT 5076  
Db 2461 CACAAGAGATAGAGTAGTATTTTGAATTTGATTTGTTTGGTCTGCTTACTGGGACTTCT 2520  
Qy 5077 TTTCTTTTCACTTAATTTTAACTTTGGTTTAAAAAGTTTTTTTATTTGGAATGTAAGTGA 5136  
Db 2521 TTTCTTTTCACTTAATTTTAACTTTGGTTTAAAAAGTTTTTTTATTTGGAATGTAAGTGA 2580  
Qy 5137 GAACCAAGAACGCACTTGAATTTTCTAAGCTCCTTAATTTGAAATGCTGTGGTTGTGTG 5196  
Db 2581 GAACCAAGAACGCACTTGAATTTTCTAAGCTCCTTAATTTGAAATGCTGTGGTTGTGTG 2640  
Qy 5197 TTTTGTCTTTCTTTAAATAAAACGATGATGATTAATTTAAGTGAAGC 5240  
Db 2641 TTTTGTCTTTCTTTAAATAAAACGATGATGATTAATTTAAGTGAAGC 2684

## RESULT 11

ABK35706

ID ABK35706 standard; cDNA; 1548 BP.

XX

AC ABK35706;

DT 08-MAY-2002 (first entry)

XX

DE cDNA sequence #97 encoding novel human secreted protein.

XX

KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
immune deficiency disorder; blood disorder; inflammatory disorder;  
infectious disorder; allergic condition; neurodegenerative disorder;  
liver fibrosis; coagulation disorder; gene therapy; antimicrobial;  
tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.

OS Homo sapiens.

XX

PN WO200177289-A2.

XX	18-OCT-2001.
PD	
XX	29-MAR-2001; 2001WO-US010232.
FF	
XX	06-APR-2000; 2000US-0195605P.
PR	
XX	(GEMV ) GENETICS INST INC.
XX	Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI	Merberg D, Ireacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;
PI	Clark HF, Fechole K, Howes SH, Resnick RJ, Gulukota K, Graham JR;
XX	WP1; 2002-179322/23.
DR	
XX	Six hundred and twenty three polynucleotides derived from a variety of
PT	human tissue sources which encode secreted proteins, useful for treating
PT	immune deficiencies and disorders such as autoimmune disorders.
XX	Claim 1; Page 126; 393pp; English.
PS	The present invention relates to the isolation of novel cDNA sequences
XX	which encode human secreted proteins. The cDNA sequences have been
CC	derived from a variety of human tissues. The invention also provides a
CC	method for producing proteins from these polynucleotide sequences. The
CC	proteins are useful for identifying compounds that modulate their
CC	activity and production. The sequences of the invention are useful for
CC	treating diseases such as hyperproliferative disorders (e.g. cancer),
CC	immune deficiency disorders (e.g. severe combined immunodeficiency
CC	(SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
CC	(e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis),
CC	infectious disorders (e.g. hepatitis), allergic conditions (e.g. asthma),
CC	neurodegenerative disorders (e.g. Alzheimer's disease), liver fibrosis,
CC	coagulation disorders (e.g. haemophilia), and tumours. The polynucleotide
CC	sequences of the invention are also useful in gene therapy. ABK35610-
CC	AKB36232 represent the cDNA sequences of the invention that encode for
CC	novel human secreted proteins
XX	
SQ	Sequence 1548 BP; 466 A; 273 C; 268 G; 541 T; 0 U; 0 Other;
	Query Match 27.3%; Score 1497; DB 6; Length 1548;
	Best Local Similarity 99.9%; Pred No. 0;
	Matches 1547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY	2489 AGAAATTGACCAAGCAGATATATAGTGTATTTACTCAGCAGCCACTGGAGGAAGAATGGA 2548
Db	1 AGAAATTGACCAAGCAGATATATAGTGTATTTACTCAGCAGCCACTGGAGGAAGAATGGA 60
QY	2549 TTCAAATCTTTTCATGAATGAAACAGAGCTTACTTATTTCTTGTAACAACAGGCCTTC 2608
Db	61 TTCAAATCTTTTCATGAATGAAACAGAGCTTACTTATTTCTTGTAACAACAGGCCTTC 120
QY	2609 TCTAGTAGCACCATGAGCCCTTTGGAAAACACAGATGTATATACATAGCAAAAGTTTCATTT 2668
Db	121 TCTAGTAGCACCATGAGCCCTTTGGAAAACACAGATGTATATACATAGCAAAAGTTTCATTT 180
QY	2669 CTTTACCCTTGAAACCGTGAAGTAAATCAGTGAGATCAGTGTGCTCTTGCTTTTAAATTTT 2728
Db	181 CTTTACCCTTGAAACCGTGAAGTAAATCAGTGAGATCAGTGTGCTCTTGCTTTTAAATTTT 240
QY	2729 TTTCAAGTTCAGATTTTTTAAGTTTTTGGTTTCATCATCTCTTTTAAAAAGTGTGGTTC 2788
Db	241 TTTCAAGTTCAGATTTTTTAAGTTTTTGGTTTCATCATCTCTTTTAAAAAGTGTGGTTC 300
QY	2789 CATCAAACTTGTCCAGACTTATATTTTCTAAAAACCTGGAGACAAACACATAAATACAA 2848
Db	301 CATCAAACTTGTCCAGACTTATATTTTCTAAAAACCTGGAGACAAACACATAAATACAA 360
QY	2849 AACAGTCTGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGTATCTTATAGCTTTTT 2908
Db	361 AACAGTCTGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGTATCTTATAGCTTTTT 420
QY	2909 CACAAGCCAGAACATAAATGCTGACGATGATTAATGACAGTGACTATGATCCGTGGCTCC 2968



1501 CAGCAATTGTCATAAGAAATATGATGACAGAGAAAGATTTCCTCTTC 1548

RESULT 12

ID ABZ35938 standard; cDNA; 2481 BP.

AC ABZ35938;

10-FEB-2003 (first entry)

Human secretory polynucleotide SPTM SEQ ID NO 102.

Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia; asthma; Crohn's disease; neurological disorder; epilepsy; cancer; Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease; multiple sclerosis; Parkinson's disease; cell proliferative disorder; anti-inflammatory; immunosuppressive; neuroprotective; neurotropic; neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic; antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus; secretory polynucleotide; secretory protein; gene; ss.

Homo sapiens.

WO200283876-A2.

24-OCT-2002.

27-MAR-2002; 2002WO-US009921.

29-MAR-2001; 2001US-0280067P.

29-MAR-2001; 2001US-0280069P.

16-MAY-2001; 2001US-0291280P.

17-MAY-2001; 2001US-0291829P.

17-MAY-2001; 2001US-0291849P.

19-JUN-2001; 2001US-0299428P.

20-JUN-2001; 2001US-0299776P.

20-JUN-2001; 2001US-0300010P.

(INCY) INCYTE GENOMICS INC.

Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

Duffour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshey SR;

Daughtery SC, Dam TC, Nguyen DA, Kleefeld Y, Gerstin EH;

Peralta CH, David ME, Lewis SA, Chen AJ, Panzer SR, Harris B;

Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

WPI; 2003-075543/07.

P-PSDB; ABP75491.

New human secretory proteins and polynucleotides, useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. AIDS), neurological disorders (e.g. Alzheimer's), or cell proliferations or cancers.

Claim 1; SEQ ID NO 102; 458pp + Sequence Listing; English.

The invention relates to a secretory polynucleotide (designated sptm) comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a naturally occurring polynucleotide sequence at least 90 % identical to the polynucleotide sequence, a polynucleotide complementary to them or an RNA equivalent of them. The polypeptide or polynucleotide are useful for treating, preventing or diagnosing a disease or condition associated with the expression of functional SPTM. These are particularly useful for diagnosing, treating or preventing autoimmune/inflammatory disorders (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's disease), neurological disorders (e.g. epilepsy, Huntington's disease, dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease, multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety, schizophrenia or amnesia), or cell proliferative disorders (e.g. psoriasis, polycythemia vera, or cancers including adenocarcinoma, leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain, breast, cervix or prostate). Note: The sequence data for this patent did

CC not form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 2481 BP; 662 A; 452 C; 477 G; 889 T; 0 U; 1 Other;

Query Match 24.7%; Score 1350; DB 7; Length 2481;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	97	AGGTTTATTCAGAAACATGTCACCTGCAATAGGAGGAGTAGGAGTTGGAGACAGACCA	156
DB	1081	AGGTTTATTCAGAAACATGTCACCTGCAATAGGAGGAGTAGGAGTTGGAGACAGACCA	1140
QY	157	GAACACTTCTTACTGAAGAATTACTTAATTAATGCAACCAAAAGAGTAGTGTTCAGG	216
DB	1141	GAACACTTCTTACTGAAGAATTACTTAATTAATGCAACCAAAAGAGTAGTGTTCAGG	1200
QY	217	AAATTCCTTTTCCACTATTTTATTTTATTTGGTTAATTAATTAATGATGATGATCCAA	276
DB	1201	AAATTCCTTTTCCACTATTTTATTTTATTTGGTTAATTAATTAATGATGATGATCCAA	1260
QY	277	ATAAGAAATATGAAGAAGTGCCTAATATAGAACTCAATCCTATGGCAAGTTTACTCTTT	336
DB	1261	ATAAGAAATATGAAGAAGTGCCTAATATAGAACTCAATCCTATGGCAAGTTTACTCTTT	1320
QY	337	CTAATCTAATCTTTGGATATATCTCCAGTGCACTAATTAATTAATGCAACATGATGAGAAAG	396
DB	1321	CTAATCTAATCTTTGGATATATCTCCAGTGCACTAATTAATTAATGCAACATGATGAGAAAG	1380
QY	397	TGCTACTGATCATCTACCTGATGTCATTAATTAATTAATGCAACATGATGAGAAAGAA	456
DB	1381	TGCTACTGATCATCTACCTGATGTCATTAATTAATTAATGCAACATGATGAGAAAGAA	1440
QY	457	TGTTAATCCTCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGGTTTCAAGATCCCA	516
DB	1441	TGTTAATCCTCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGGTTTCAAGATCCCA	1500
QY	517	TGTCCTATGAACCTTCGTTTTTTCCTGATATGATCCAGTATCTTCTATTTATATGATT	576
DB	1501	TGTCCTATGAACCTTCGTTTTTTCCTGATATGATCCAGTATCTTCTATTTATATGATT	1560
QY	577	CAAGAGCTGGCTGTTCATAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTCCAG	636
DB	1561	CAAGAGCTGGCTGTTCATAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTCCAG	1620
QY	637	TTTACAAGCATCCATAGATGCTGCCATTAACAGTTGAAGCAATGTTCTCTTTGGA	696
DB	1621	TTTACAAGCATCCATAGATGCTGCCATTAACAGTTGAAGCAATGTTCTCTTTGGA	1680
QY	697	AGGAGCTGGAGTCACTAAAGCTGTTTATATGGAGAACTGCTGTTGAGAAATAGATA	756
DB	1681	AGGAGCTGGAGTCACTAAAGCTGTTTATATGGAGAACTGCTGTTGAGAAATAGATA	1740
QY	757	CTTTCCCGGAGGAGTAATTTTAAATACCTAGTTATAGCATTTTCACCTTTGGATAT	816
DB	1741	CTTTCCCGGAGGAGTAATTTTAAATACCTAGTTATAGCATTTTCACCTTTGGATAT	1800
QY	817	TTTGGCAATTCATATCGTAGCAGAAAGAAAAAATAAAGAAATTTTAAAGATAA	876
DB	1801	TTTGGCAATTCATATCGTAGCAGAAAGAAAAAATAAAGAAATTTTAAAGATAA	1860
QY	877	TGGGACTTCATGATGCTCTTTTGGCTTCTGGGTTCTTCTATATACAAGTTAAATTT	936
DB	1861	TGGGACTTCATGATGCTCTTTTGGCTTCTGGGTTCTTCTATATACAAGTTAAATTT	1920
QY	937	TTCTTATGCTCTTCTTATGGCAGTCATTCGACAGCTTCTTTGTTATTTCTCAAGTA	996
DB	1921	TTCTTATGCTCTTCTTATGGCAGTCATTCGACAGCTTCTTTGTTATTTCTCAAGTA	1980
QY	997	GCAGCATTTGTGATTTCTGCTTTTTTCTTTTCTTTATGATATCATCTGTATTTTGTCT	1056
DB	1981	GCAGCATTTGTGATTTCTGCTTTTTTCTTTTCTTTATGATATCATCTGTATTTTGTCT	2040



Db 721 ATGTTTGAAGCAAGAAAATGATTGGCAATTTCTCCACAGTTAGATATACACTTTTGATCT 780  
 QY 1824 TTGACAGTAGAAGAAAATTTATCAATTTTGGCTTCTCATCAAGAGGATACCAAGCAACAT 1883  
 Db 781 TTGACAGTAGAAGAAAATTTATCAATTTTGGCTTCTCATCAAGAGGATACCAAGCAACAT 840  
 QY 1884 ATAATAACAAGAGTGCAGAAAGGTTTACTAGATTTAGACATGCAGACTATCAAGATAAC 1943  
 Db 841 ATAATAACAAGAGTGCAGAAAGGTTTACTAGATTTAGACATGCAGACTATCAAGATAAC 900  
 QY 1944 CAGCTAAAAATTAAGTGGTGCCTAAAAAGAAAGCTGTCTATTAGGAATTTGCTTTCTT 2003  
 Db 901 CAGCTAAAAATTAAGTGGTGCCTAAAAAGAAAGCTGTCTATTAGGAATTTGCTTTCTT 960  
 QY 2004 GGGACCCCAAGATAGTCTGCTGTAGATGAACCAACAGCTGGAATCGACCCCTGTTCTCGA 2063  
 Db 961 GGGACCCCAAGATAGTCTGCTGTAGATGAACCAACAGCTGGAATCGACCCCTGTTCTCGA 1020  
 QY 2064 CATATTGTATGAATCTTTTAAAAATACAGAAAGCAATCGGGTGACAGTGTTCAGTACT 2123  
 Db 1021 CATATTGTATGAATCTTTTAAAAATACAGAAAGCAATCGGGTGACAGTGTTCAGTACT 1080  
 QY 2124 CATTTATGATGAGTGCAGATCTTCTGAGATGAACCAACAGCTGGAATCGATATCAAGGAATG 2183  
 Db 1081 CATTTATGATGAGTGCAGATCTTCTGAGATGAACCAACAGCTGGAATCGATATCAAGGAATG 1140  
 QY 2184 CTGAATGTGTTGGTCTTCTCAATGTTCTCAAAAGTAATGGGGATCGGCTACCGCCTG 2243  
 Db 1141 CTGAATGTGTTGGTCTTCTCAATGTTCTCAAAAGTAATGGGGATCGGCTACCGCCTG 1200  
 QY 2244 AGCATGTATACAGCAAAATTTGTCACAGATCTTCTTCTACTGGTTAAACAACAT 2303  
 Db 1201 AGCATGTATACAGCAAAATTTGTCACAGATCTTCTTCTACTGGTTAAACAACAT 1260  
 QY 2304 ATACCTGGAGTACTTTTATCAACAGAAATGACCAACAACTTGTTAGTCTTGCTTTTC 2363  
 Db 1261 ATACCTGGAGTACTTTTATCAACAGAAATGACCAACAACTTGTTAGTCTTGCTTTTC 1320  
 QY 2364 AGGACATGGACAAATTTTCAGG 2386  
 Db 1321 AAGGACATGGACAAATTTTCAGG 1343

RESULT 14  
 ABZ35926  
 ID ABZ35926 standard; cDNA; 1632 BP.  
 XX  
 AC ABZ35926;  
 XX  
 DT 10-FEB-2003 (first entry)  
 XX  
 DE Human secretory polynucleotide SPTM SEQ ID NO 90.  
 XX  
 KW Human; SPTM; autoimmune disorder; inflammatory disorder; AIDS; anaemia;  
 KW asthma; Crohn's disease; neurological disorder; epilepsy; cancer;  
 KW Huntington's disease; Alzheimer's disease; Creutzfeldt-Jakob disease;  
 KW multiple sclerosis; Parkinson's disease; cell proliferative disorder;  
 KW anti-inflammatory; immunosuppressive; neuroprotective; nontropic;  
 KW neuroleptic; anticonvulsant; cytostatic; antiparkinsonian; anxiolytic;  
 KW antipsoriatic; antianaemic; anti-HIV; human immunodeficiency virus;  
 KW secretory polynucleotide; secretory protein; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200283876-A2.  
 XX  
 PD 24-OCT-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-US009921.  
 XX  
 PR 29-MAR-2001; 2001US-0280067P.  
 PR 29-MAR-2001; 2001US-0280068P.  
 PR 16-MAY-2001; 2001US-0291280P.

PR 17-MAY-2001; 2001US-0291829P.  
 PR 17-MAY-2001; 2001US-0291849P.  
 PR 19-JUN-2001; 2001US-0299428P.  
 PR 20-JUN-2001; 2001US-0299776P.  
 PR 20-JUN-2001; 2001US-0300001P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 PI Dufour GE, Hillman JB, Yu JY, Tuason O, Yap PE, Amshay SR;  
 PI Daugherty SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstin EH;  
 PI Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 PI Flores V, Marwaha R, Lo A, Lan RV, Urashka NE;  
 XX WPI: 2003-075543/07.  
 DR P-PSDB; ABP75479.  
 DR  
 XX  
 PT New human secretory proteins and polynucleotides, useful for diagnosing,  
 PT treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),  
 PT neurological disorders (e.g. Alzheimer's), or cell proliferations or  
 PT cancers.  
 XX  
 PS Claim 1; SEQ ID NO 90; 458pp + Sequence Listing; English.  
 XX  
 CC The invention relates to a secretory polynucleotide (designated sptm)  
 CC comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a  
 CC naturally occurring polynucleotide sequence at least 90 % identical to  
 CC the polynucleotide sequence, a polynucleotide complementary to them or an  
 CC RNA equivalent of them. The polypeptide or polynucleotide are useful for  
 CC treating, preventing or diagnosing a disease or condition associated with  
 CC the expression of functional SPTM. These are particularly useful for  
 CC diagnosing, treating or preventing autoimmune/inflammatory disorders  
 CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's  
 CC disease), neurological disorders (e.g. Creutzfeldt-Jakob disease,  
 CC dementia, stroke, Alzheimer's disease, epilepsy, Huntington's disease,  
 CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,  
 CC schizophrenia or amnesia), or cell proliferative disorders (e.g.  
 CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,  
 CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,  
 CC breast, cervix or prostate). Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 1632 BP; 527 A; 205 C; 301 G; 599 T; 0 U; 0 Other;  
 Query Match 15.1%; Score 829; DB 7; Length 1632;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 829; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4647 AGATGTATCGGAACAGTACAAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAA 4706  
 Db 218 AGATGTATCGGAACAGTACAAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAA 277  
 QY 4707 ATTAATTTGAAGACTGGATAGAAAACCTAGAGTAGACCGCTTCAAGAGAAATTCAG 4766  
 Db 278 ATTAATTTGAAGACTGGATAGAAAACCTAGAGTAGACCGCTTCAAGAGAAATTCAG 337  
 QY 4767 TATATTTTCCCAATCGAAGCCGTGAGAAAGTTTCTTCTATTTTGGCTTATAAAAT 4826  
 Db 338 TATATTTTCCCAATCGAAGCCGTGAGAAAGTTTCTTCTATTTTGGCTTATAAAAT 397  
 QY 4827 CCTAAGAGAGATGTTTCAGTCCCTTTTCACAATCTTTTTTAAAGCTGAAGAGCTAAACAT 4886  
 Db 398 CCTAAGAGAGATGTTTCAGTCCCTTTTCACAATCTTTTTTAAAGCTGAAGAGCTAAACAT 457  
 QY 4887 GCTTTTGGCAATGAAGATATAGCTTTTCTCAAGCAACATTTGACAGAGTTTGTAGAA 4946  
 Db 458 GCTTTTGGCAATGAAGATATAGCTTTTCTCAAGCAACATTTGACAGAGTTTGTAGAA 517  
 QY 4947 CTCACCTAAAGAACCAAGAGGAGGAGATAATAGTTGTGGAACTTTAAACAGCACACTTGG 5006  
 Db 518 CTCACCTAAAGAACCAAGAGGAGGAGATAATAGTTGTGGAACTTTAAACAGCACACTTGG 577

QY	5007	TGGGACCAACACACAGAGATAGAGTAGTATTTTGAATTTTGATTTGTTCCGGTCTGCTTAC	5066
Db	578	TGGGACCAACACACAGAGATAGAGTAGTATTTTGAATTTTGATTTGTTCCGGTCTGCTTAC	637
QY	5067	TGGGACTTCTTTTCTTTTCTTTTTCACCTTAATTTTAACTTTTGGTTTTAAAAAGTTTTTTTATTGGAAT	5126
Db	638	TGGGACTTCTTTTCTTTTCTTTTTCACCTTAATTTTAACTTTTGGTTTTAAAAAGTTTTTTTATTGGAAT	697
QY	5127	GTGAACGGGAGCAACAGACGCACTTGAATTTTTTCTAAGCTCCTTAATTTGAATGCTG	5186
Db	698	GTGAACGGGAGCAACAGACGCACTTGAATTTTTTCTAAGCTCCTTAATTTGAATGCTG	757
QY	5187	TGGTGTGCTGTTTGTCTTCTTTTCTTTTAAATAAAACGATGTATATTAATTAAGTGAAGCTGCATG	5246
Db	758	TGGTGTGCTGTTTGTCTTCTTTTCTTTTAAATAAAACGATGTATATTAATTAAGTGAAGCTGCATG	817
QY	5247	TTTGATTTGAAGTATATTGAACATATAGTTTCTGATGTCATCTTTTTCACCATTCAGAAA	5306
Db	818	TTTGATTTGAAGTATATTGAACATATAGTTTCTGATGTCATCTTTTTCACCATTCAGAAA	877
QY	5307	CAGTCTCTCTGAAATTTGTGATTTAAAGAAATGTAATAGAAATGTTTTATTTTAAAGTTA	5366
Db	878	CAGTCTCTCTGAAATTTGTGATTTAAAGAAATGTAATAGAAATGTTTTATTTTAAAGTTA	937
QY	5367	TCTTTTAAAGTTATGCCATCTTCTTTTAAATAAGTAGTAGTAATGTTTCCCAATCTAAATAAAAAAC	5426
Db	938	TCTTTTAAAGTTATGCCATCTTCTTTTAAATAAGTAGTAGTAATGTTTCCCAATCTAAATAAAAAAC	997
QY	5427	TAAATCAATCAATTAATGCATAGAAAAGATACATAAAGCAATGTGAAAGTT	5475
Db	998	TAAATCAATCAATTAATGCATAGAAAAGATACATAAAGCAATGTGAAAGTT	1046

RESULT 15

AAC81717

ID

AAC81717 standard; DNA; 1506 BP.

XX

AAC81717;

XX

26-FEB-2001 (first entry)

XX

Human secreted protein coding sequence SEQ ID NO: 18.

XX

Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

XX

antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;

XX

vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic;

XX

candidant; gene therapy; cancer; immune disorder; cardiovascular disorder;

XX

neurological disease; infection; human; secreted protein; ss.

XX

Homo sapiens.

XX

OS

XX

PN

XX

WO200061627-A1.

XX

19-OCT-2000.

XX

06-APR-2000; 2000WO-US009067.

XX

09-APR-1999; 99US-0128697P.

XX

20-JAN-2000; 2000US-0176929P.

XX

(HUMA-) HUMAN GENOME SCI INC.

PA

(ROSE/) ROSEN C A.

XX

Rosen Ch, Ruben SM, Komatsoulis G;

XX

WPI; 2000-647419/62.

XX

P-PSDB; AAB45393.

XX

New nucleic acid molecules encoding 49 human secreted proteins for

XX

diagnosing, preventing, treating or ameliorating medical conditions and

XX

used as food additives or preservatives.

XX

Claim 1; Page 429; 521pp; English.

XX		The invention relates to the isolation of genes AAC81710-C81758 encoding
CC	49	human secreted proteins AAB45386-B54534. The genes can be used to
CC		generate fusion proteins (SEQID) for increasing the stability of the
CC		immunoglobulin Fc portion (SEQID) for increasing the stability of the
CC		fusion protein as compared to the human protein only. The genes and
CC		proteins are useful for preventing, ameliorating or treating medical
CC		conditions, e.g. by protein or gene therapy. The genes are isolated from
CC		a range of human tissues disclosed in the specification. The nucleic
CC		acids, proteins, antibodies and (ant)agonists are useful in the
CC		diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC		ovarian cancer, and other cancers of the adrenal gland, bone, bone
CC		marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
CC		immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC		anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC		multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC		cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC		; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC		infectious diseases such as viral, bacterial, fungal and parasitic
CC		infections
XX		
SQ		Sequence 1506 BP; 509 A; 218 C; 290 G; 485 T; 0 U; 4 Other;
		Query Match 14.3%; Score 783; DB 3; Length 1506;
		Best Local Similarity 100.0%; Pred. No. 0;
		Matches 783; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	4648	GATGTATCGGAACAGTACAACATCTAAAGAGTAATAATTGGAAAAGGCTACTTTTGAAA 4707
Db	485	GATGTATCGGAACAGTACAACATCTAAAGAGTAATAATTGGAAAAGGCTACTTTTGAAA 544
Qy	4708	TTAATTTGAAGGACTGGATAGAAAAAACCCTAGAAGTAGACCGCCTTCAAAGAGAAATTCAGT 4767
Db	545	TTAATTTGAAGGACTGGATAGAAAAAACCCTAGAAGTAGACCGCCTTCAAAGAGAAATTCAGT 604
Qy	4768	ATATTTTCCCNAATGCNAGCCGTCAGRAAGTTTTTCTCTATTTTGGCTTTATAAAATTC 4827
Db	605	ATATTTTCCCNAATGCNAGCCGTCAGRAAGTTTTTCTCTATTTTGGCTTTATAAAATTC 664
Qy	4828	CTAAGGAAGATGTTCAAGTCCCTTTTCCAATCTTTTTTAAGCTGGAGAAGCTAAACATG 4897
Db	665	CTAAGGAAGATGTTCAAGTCCCTTTTCCAATCTTTTTTAAGCTGGAGAAGCTAAACATG 724
Qy	4888	CTTTTGGCATTTGAAGATATAGCTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAAC 4947
Db	725	CTTTTGGCATTTGAAGATATAGCTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAAC 784
Qy	4948	TCACTAAGAACCAAGAGGAGGAAGATAATAGTTGTGGAATTTTAAACAGCACACTTTGGT 5007
Db	785	TCACTAAGAACCAAGAGGAGGAAGATAATAGTTGTGGAATTTTAAACAGCACACTTTGGT 844
Qy	5008	GGGAACGACACAGAGATAGTAGTATTTTGAATTTCTATTGTTTCGGTCTGCTTACT 5067
Db	845	GGGAACGACACAGAGATAGTAGTATTTTGAATTTCTATTGTTTCGGTCTGCTTACT 904
Qy	5068	GGGACTTCTTTCTTTTCACTTTAATTTTAACCTTTGGTTTAAAAAGTTTTTATTGGAATG 5127
Db	905	GGGACTTCTTTCTTTTCACTTTAATTTTAACCTTTGGTTTAAAAAGTTTTTATTGGAATG 964
Qy	5128	GTAACCTGGGAACCAAGAACCGCACTTGAATTTTTTCTTAAGCTCCTTAATTTGMAATGCTGT 5187
Db	965	GTAACCTGGGAACCAAGAACCGCACTTGAATTTTTTCTTAAGCTCCTTAATTTGMAATGCTGT 1024
Qy	5188	GGTTGTGTGTTTGTCTTTCTTTTAAATPAAACGTA TGTAATAA TAAGTGAAGCTGCATGT 5247
Db	1025	GGTTGTGTGTTTGTCTTTCTTTTAAATPAAACGTA TGTAATAA TAAGTGAAGCTGCATGT 1084
Qy	5248	TTGTATTGAAGTATATTGAACACTATAGTTTGTATGTCATCTTTTTCACCAATTCAGAAAC 5307
Db	1085	TTGTATTGAAGTATATTGAACACTATAGTTTGTATGTCATCTTTTTCACCAATTCAGAAAC 1144
Qy	5308	AGTGCTTCTGAA TTGTGATTTAAAGAAATGTGAA TAGAATAGT TTTTATTTAAGTTAT 5367

Db 1145 AGTGCTTCTGAATTTGTGATTTAAAGGAATTGTAATAGATAGTTTATTTTTTAAGTTAT 1204  
QY 5368 CTTTAAAGTTTATGCCATCTCTTAAATAAGTACGTAAATGTTCCAAATCTAAATAAAAACT 5427  
Db 1205 CTTTAAAGTTTATGCCATCTCTTAAATAAGTACGTAAATGTTCCAAATCTAAATAAAAACT 1264  
QY 5428 AAT 5430  
Db 1265 AAT 1267

Search completed: April 12, 2004, 11:49:10  
Job time : 1309 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2004, 21:18:29 ; Search time 13649 Seconds  
(without alignments)  
17386.117 Million cell up

Title: US-10-090-458-4  
Perfect score: 5475  
Sequence: 1 gcgtcgcgcgcctcgacag.....cataaagcaatgtgaaagt 5475

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

```
Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	5474	100.0	5475	6	AX537473	AX537473 Sequence	
2	5373.2	98.1	6525	6	AX473847	AX473847 Sequence	
3	5373.2	98.1	6525	9	AY028897	AY028897 Homo sapi	
4	5366.6	97.8	5463	6	AX537470	AX537470 Sequence	
5	5335.6	93.8	6369	6	AX392931	AX392931 Sequence	
6	5083.2	92.8	5096	9	HS4275973	AX275973 Homo sapi	
7	5014.6	91.6	5262	6	AX417828	AX417828 Sequence	
8	4922.6	89.9	4929	6	AX417824	AX417824 Sequence	
9	4811.6	87.9	4917	6	AX537472	AX537472 Sequence	
10	4759.8	86.9	4785	6	AX417826	AX417826 Sequence	
11	3891.7	71.1	5347	10	AF491842	AF491842 Mus muscu	
12	3879.8	70.9	5170	10	RN0436052	AJ426052 Rattus n	
13	3873	70.7	4929	10	AB097675	AB097675 Mus muscu	
14	3820.4	69.8	4945	10	RN0550165	AJ550165 Rattus n	
15	3812.6	69.6	4078	9	HSN808823	AX830336 Sequence	
16	3256	59.5	3347	6	AX833036	AX094416 Homo sapi	
17	3256	59.5	3347	6	AX094416	AX714300 Sequence	
18	3175.8	58.0	3268	6	AX714300	AX056533 Homo sapi	
19	3175.8	58.0	3268	9	AX056533	AJ512612 Homo sapi	
20	2815.2	51.4	2845	9	HS4513612	AB067475 Homo sapi	
21	2646.4	48.3	3112	9	AB067475	AX647447 Homo sapi	
22	2436.2	44.5	5243	9	HSN807592	AX122803 Homo sapi	
23	1807.6	33.0	1364	9	AK122803	AK1780218 Sequence	
24	1668.4	30.5	2325	6	AX780218	AK058170 Homo sapi	
25	1560	28.5	1560	9	AK058170	AK129463 Mus muscu	
26	1484.4	27.1	4921	10	AK129463	AX331591 Sequence	
27	1182	21.6	1346	6	AX331591	AX332287 Sequence	
28	1182	21.6	1346	6	AX332287	AX410760 Sequence	
29	1182	21.6	1346	6	AX410760	AX473850 Sequence	
30	1149.8	21.0	6181	6	AX473850	AX57703 Sequence	
31	1149.8	21.0	6181	9	AY028900	AX247065 Homo sapi	
32	1149.2	21.0	5722	6	AX57703	AF423307 Homo sapi	
33	1149.2	21.0	6333	9	AF423307	AY028899 Homo sapi	
34	1126.2	20.6	4875	9	AF423307	AX535879 Sequence	
35	1121.4	20.5	5981	6	AX473849	AX535877 Sequence	
36	1121.4	20.5	5981	6	AY028899	AX535877 Sequence	
37	1119.8	20.5	4875	6	AX535879	AX473848 Sequence	
38	1119.8	20.5	5018	6	AX535877	AX028898 Homo sapi	
39	1090.6	19.9	5296	6	AX473848	AF373290 Homo sapi	
40	1090.6	19.9	5296	9	AX028898	AX575480 Sequence	
41	1089	19.9	4984	9	AF373290	AX538009 Sequence	
42	1089	19.9	5211	6	AX575480	AX538007 Sequence	
43	1085.8	19.8	4854	6	AX538009	AX657701 Sequence	
44	1085.8	19.8	5332	6	AX538007		
45	1084	19.8	5680	6	AX657701		

## ALIGNMENTS

RESULT	1
AX537473	
LOCUS	DNA linear PAT 23-NOV-2002
DEFINITION	Sequence 4 from Patent WO02070690.
ACCESSION	AX537473
VERSION	AX537473.1 GI:25269282
KEYWORDS	Homo sapiens (human)
SOURCE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. .1
REFERENCE	Chen,H., Kilinski,L. and le Bihan,S.
AUTHORS	
TITLE	Apcas transporter and uses thereof
JOURNAL	Patent: WO 02070690-A 4 12-SEP-2002;

FEATURES		Active Pass Pharmaceuticals, Inc. (CA)	
source	Location/Qualifiers		
	1..5475		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match		100.0%; Score 5474; DB 5; Length 5475;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 5475; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GGGTCGGGCGCCCTCGCACAGATCCNAGCTGGTGCACCCGCACTGAGTCAACAGACTCCAGC	60
DB	1	GGGTCGGGCGCCCTCGCACAGATCCNAGCTGGTGCACCCGCACTGAGTCAACAGACTCCAGC	60
QY	61	GGGTCGAGGCTGACAGCTCTGCGGCTCGGCGCTCGGCGCTGAGGTTTATTCAGAAAACATGTC	120
DB	61	GGGTCGAGGCTGACAGCTCTGCGGCTCGGCGCTCGGCGCTGAGGTTTATTCAGAAAACATGTC	120
QY	121	CTGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGAACACTTCTACTGAAGAAATTACT	180
DB	121	CTGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGAACACTTCTACTGAAGAAATTACT	180
QY	181	TAAATTAATGCAAGACCAAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTATTTTTT	240
DB	181	TAAATTAATGCAAGACCAAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTATTTTTT	240
QY	241	TATTTTGGTAAATTAATTAATGATGATGCATCCAAATAGAAATATGAAGAGTGCTA	300
DB	241	TATTTTGGTAAATTAATTAATGATGATGCATCCAAATAGAAATATGAAGAGTGCTA	300
QY	301	ATATAGAACTCAATCCTATGGACAAAGTTTACTTTTCTAATCTAATCTTCTGGATATATC	360
DB	301	ATATAGAACTCAATCCTATGGACAAAGTTTACTTTTCTAATCTAATCTTCTGGATATATC	360
QY	361	CAGTGAATTAATTAACAGCAGCATCATGCAAGAGTGTCTACTGATCACTACTGATG	420
DB	361	CAGTGAATTAATTAACAGCAGCATCATGCAAGAGTGTCTACTGATCACTACTGATG	420
QY	421	TCAATTAATTAATGAGAAATATCAAAATGAAAGAAATGTTAACTCAAGTCTCTCTAAGC	480
DB	421	TCAATTAATTAATGAGAAATATCAAAATGAAAGAAATGTTAACTCAAGTCTCTCTAAGC	480
QY	481	CGAGCAACTTTGTAGGTGGTTTCAAGAGCTCCATGCTCTATGAACTTGGTTTTTTTC	540
DB	481	CGAGCAACTTTGTAGGTGGTTTCAAGAGCTCCATGCTCTATGAACTTGGTTTTTTTC	540
QY	541	CTGATATGATCCAGTATCTTCTATTTATATGGATTCAAGAGCTGGCTGTTCAAAATCAT	600
DB	541	CTGATATGATCCAGTATCTTCTATTTATATGGATTCAAGAGCTGGCTGTTCAAAATCAT	600
QY	601	GTGAGGCTGCTCAGTACTGGTCTCAGGTTTCAAGTTTCAAGCATCCATAGATGCTG	660
DB	601	GTGAGGCTGCTCAGTACTGGTCTCAGGTTTCAAGTTTCAAGCATCCATAGATGCTG	660
QY	661	CCATTATACAGTGAAGCCAAATGTTCTCTTGGAGGAGCTGGAGTCAACTAAAGCTG	720
DB	661	CCATTATACAGTGAAGCCAAATGTTCTCTTGGAGGAGCTGGAGTCAACTAAAGCTG	720
QY	721	TTATTATGGGAAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTAATTTAA	780
DB	721	TTATTATGGGAAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTAATTTAA	780
QY	781	TATACCTAGTTATAGCAATTTTACCTTTTGGATCTTTTGGCAATTCATATCCTAGCAG	840
DB	781	TATACCTAGTTATAGCAATTTTACCTTTTGGATCTTTTGGCAATTCATATCCTAGCAG	840
QY	841	AAAAAGAAAAAATAAAGAAATTTTAAAGATAATCGGACTTCAATGATCTGCTTTT	900
DB	841	AAAAAGAAAAAATAAAGAAATTTTAAAGATAATCGGACTTCAATGATCTGCTTTT	900
QY	901	GGCTTTCTGGGTTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCCTCTTATGGCAG	960

DB	901	GGCTTTCTGGGTTCTTCTATATACAGTTTAAATTTTCTTAATGTCCTTCTTATGGCAG	960
QY	961	TCATTGCGACAGCTCTTTTGTATTTTCTCAAAGTAGCAGCATTTGATATTTTCTGCTTT	1020
DB	961	TCATTGCGACAGCTCTTTTGTATTTTCTCAAAGTAGCAGCATTTGATATTTTCTGCTTT	1020
QY	1021	TTTTCTCTTATGATATCATCTGTATTTTGTCTTAAATGCTGACACCTCTTTTAAAA	1080
DB	1021	TTTTCTCTTATGATATCATCTGTATTTTGTCTTAAATGCTGACACCTCTTTTAAAA	1080
QY	1081	AATCAAAAACATGTGGGAATAGTGAATTTTGTGTCTGTGGCTTTTGGATTTATGGCC	1140
DB	1081	AATCAAAAACATGTGGGAATAGTGAATTTTGTGTCTGTGGCTTTTGGATTTATGGCC	1140
QY	1141	TTATGATAATCTCATAGAAAGTTTCCCAATCTGTAGTGTGGCTTTTCACTGCTTCT	1200
DB	1141	TTATGATAATCTCATAGAAAGTTTCCCAATCTGTAGTGTGGCTTTTCACTGCTTCT	1200
QY	1201	GTCACTGTACTTTTGTGATTGATTGACAGGTCATGCAATTTAGAAGATTTTAAATGAG	1260
DB	1201	GTCACTGTACTTTTGTGATTGATTGACAGGTCATGCAATTTAGAAGATTTTAAATGAG	1260
QY	1261	GTGCTTCATTTTCAAAATTTGACTGCGGCCCATATCTCTAAATTTTCAATATCATGC	1320
DB	1261	GTGCTTCATTTTCAAAATTTGACTGCGGCCCATATCTCTAAATTTTCAATATCATGC	1320
QY	1321	TCACACTTAATAGTATATTTCTATGCTCTGCTGCTCTATCTGATCAAGTCATTCAG	1380
DB	1321	TCACACTTAATAGTATATTTCTATGCTCTGCTGCTCTATCTGATCAAGTCATTCAG	1380
QY	1381	GGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTTGGTCA	1440
DB	1381	GGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTTGGTCA	1440
QY	1441	GCAAAAGAAATTTATGAGGAGTTTATCAGAGGCAATGTTAAATGCAATATTTAGTTG	1500
DB	1441	GCAAAAGAAATTTATGAGGAGTTTATCAGAGGCAATGTTAAATGCAATATTTAGTTG	1500
QY	1501	AAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATTAAGATTAGTGA	1560
DB	1501	AAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATTAAGATTAGTGA	1560
QY	1561	TTCAAGAGCATATCAGAAAGAGGTTGAAAATGTCGAGGCTTTGAGAAATTTGCTATTG	1620
DB	1561	TTCAAGAGCATATCAGAAAGAGGTTGAAAATGTCGAGGCTTTGAGAAATTTGCTATTG	1620
QY	1621	ACATATATGAGGCTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGAAAGATACAT	1680
DB	1621	ACATATATGAGGCTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGAAAGATACAT	1680
QY	1681	TGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATATGGAC	1740
DB	1681	TGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATATGGAC	1740
QY	1741	ACAGAGCTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAATGATTTGGCATTTTCCAC	1800
DB	1741	ACAGAGCTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAATGATTTGGCATTTTCCAC	1800
QY	1801	AGTTAGATATACATTTGATGTTTGAAGTGAAGAAAATTTTCAATTTTGGCTTCAA	1860
DB	1801	AGTTAGATATACATTTGATGTTTGAAGTGAAGAAAATTTTCAATTTTGGCTTCAA	1860
QY	1861	TCAAAGGATATACAGCAACCAATATAATACAGAGAGTGCAGAGGTTTACTTAGATTAG	1920
DB	1861	TCAAAGGATATACAGCAACCAATATAATACAGAGAGTGCAGAGGTTTACTTAGATTAG	1920
QY	1921	ACATGAGACTATCAAGATAACCAAGCTTAAATAATTAAGTGGTGTCAAAAAGAAAGC	1980
DB	1921	ACATGAGACTATCAAGATAACCAAGCTTAAATAATTAAGTGGTGTCAAAAAGAAAGC	1980
QY	1981	TGTCATTAGGAATTTGCTTTCTTGGGAACCAAGATATCTGCTAGATGAACCAACAG	2040



Db 1981 TGTCAATTAGGAATTGCTGTTCTTGGAAACCAAGATAGTCTGCTAGATGAACCAACAG 2040  
Qy 2041 CTGGAAATGAGACCCCTGTTCTCGACATATGTAATGAATCTTTTAAATACAGAAAGCCA 2100  
Db 2041 CTGGAAATGAGACCCCTGTTCTCGACATATGTAATGAATCTTTTAAATACAGAAAGCCA 2100  
Qy 2101 ATCGGGTGACAGTGTTCAGTACTCAATTCATGGATGAAGCTGACATCTTTCAGATAGGA 2160  
Db 2101 ATCGGGTGACAGTGTTCAGTACTCAATTCATGGATGAAGCTGACATCTTTCAGATAGGA 2160  
Qy 2161 AAGCTGTGATATCAACAGGAATGCTGAATGTTGTTCTTCAATGTTCTCTCAAAAGTA 2220  
Db 2161 AAGCTGTGATATCAACAGGAATGCTGAATGTTGTTCTTCAATGTTCTCTCAAAAGTA 2220  
Qy 2221 AATGGGGATCGGCTACCGCTGAGCATGTATAGACAAATATTTGGCCACAGAAATCTC 2280  
Db 2221 AATGGGGATCGGCTACCGCTGAGCATGTATAGACAAATATTTGGCCACAGAAATCTC 2280  
Qy 2281 TTTCTTCACTGGTTAAACAAATATACCTGGAGCTACTTTTATCAACAGAAATGACCAAC 2340  
Db 2281 TTTCTTCACTGGTTAAACAAATATACCTGGAGCTACTTTTATCAACAGAAATGACCAAC 2340  
Qy 2341 AACTTGTGTATAGTCTTTCAGGACATGGACAAATTTTCAGGTTTGTCTGCCC 2400  
Db 2341 AACTTGTGTATAGTCTTTCAGGACATGGACAAATTTTCAGGTTTGTCTGCCC 2400  
Qy 2401 TAGACAGTCAATCAAAATTTGGTGTCTATTTATGTTGTTTCCATGACGACTTTGGAAG 2460  
Db 2401 TAGACAGTCAATCAAAATTTGGTGTCTATTTATGTTGTTTCCATGACGACTTTGGAAG 2460  
Qy 2461 ACGTATTTTAAAGCTAGAAGTTGAAGCAGAAATTTGACCAAGCAGATATATAGTGTATTA 2520  
Db 2461 ACGTATTTTAAAGCTAGAAGTTGAAGCAGAAATTTGACCAAGCAGATATATAGTGTATTA 2520  
Qy 2521 CTGACAGCCATCGAGGAAGAAATGGATCAAAATCTTTTGTGATGAATGAACAGAGCT 2580  
Db 2521 CTGACAGCCATCGAGGAAGAAATGGATCAAAATCTTTTGTGATGAATGAACAGAGCT 2580  
Qy 2581 TACTTATTTCTTGAACCAAGGCTTCTAGTGAGCACCATTGAGCCTTTTGGAAACAAAC 2640  
Db 2581 TACTTATTTCTTGAACCAAGGCTTCTAGTGAGCACCATTGAGCCTTTTGGAAACAAAC 2640  
Qy 2641 AGATGTATACATAGCAAAAGTTTCAATTTTACCTTGAACCGTGAAGTAAATCAGTGA 2700  
Db 2641 AGATGTATACATAGCAAAAGTTTCAATTTTACCTTGAACCGTGAAGTAAATCAGTGA 2700  
Qy 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAATTTTATGTTTGGTTC 2760  
Db 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAATTTTATGTTTGGTTC 2760  
Qy 2761 ATCACTCTTTTAAATGCTGGTTCGCATCAAACTTTGTTCCAGACTTATATTTCTAA 2820  
Db 2761 ATCACTCTTTTAAATGCTGGTTCGCATCAAACTTTGTTCCAGACTTATATTTCTAA 2820  
Qy 2821 AACCTGGAGACAAACCAATAAATCAAAACAAAGTCTGCTTCTTCAAAATTTCTGCACT 2880  
Db 2821 AACCTGGAGACAAACCAATAAATCAAAACAAAGTCTGCTTCTTCAAAATTTCTGCACT 2880  
Qy 2881 CAGATATCAGTCACTTATAGCTTTTTCACAGCCAGAACATATGTTGACGATGATTA 2940  
Db 2881 CAGATATCAGTCACTTATAGCTTTTTCACAGCCAGAACATATGTTGACGATGATTA 2940  
Qy 2941 ATGACAGTGACTATGATCCGTTGGCTCCCATAGTGGGCTTTAAATGTGATGATTCAG 3000  
Db 2941 ATGACAGTGACTATGATCCGTTGGCTCCCATAGTGGGCTTTAAATGTGATGATTCAG 3000  
Qy 3001 AAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTATGTTTATTTCTTCTATAT 3060  
Db 3001 AAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTATGTTTATTTCTTCTATAT 3060  
Qy 3061 TAGTGAATATCATAGTAACTACTATCTTTTATCAATTTAAATGTGACTGAAACCATCCAGA 3120  
Db 3061 TAGTGAATATCATAGTAACTACTATCTTTTATCAATTTAAATGTGACTGAAACCATCCAGA 3120

Qy 3121 TCTGGAGTACCCCATTTCTTTCAAGAAATPACTGATATAGTCTTTTAAATTTGAGCTGATTT 3180  
Db 3121 TCTGGAGTACCCCATTTCTTTCAAGAAATPACTGATATAGTCTTTTAAATTTGAGCTGATTT 3180  
Qy 3181 TTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCACTTACTTTGGCATGGA 3240  
Db 3181 TTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCACTTACTTTGGCATGGA 3240  
Qy 3241 ATGCAGAGAAATCATAGATCAAAAGCTTATACCTCAACTTAAACTTTTCAGGCTCTTTGCCAT 3300  
Db 3241 ATGCAGAGAAATCATAGATCAAAAGCTTATACCTCAACTTAAACTTTTCAGGCTCTTTGCCAT 3300  
Qy 3301 CTGCATATTTGGATTTGGAACAAAGCTGTTGTTGATATCCCTTATTTTTATCATTTCTTATTT 3360  
Db 3301 CTGCATATTTGGATTTGGAACAAAGCTGTTGTTGATATCCCTTATTTTTATCATTTCTTATTT 3360  
Qy 3361 TGNATGCTAGGAGCTTTATTTGGCAATTTCAATTAAGGATATATTTTTATCTGTAAAGTTCC 3420  
Db 3361 TGNATGCTAGGAGCTTTATTTGGCAATTTCAATTAAGGATATATTTTTATCTGTAAAGTTCC 3420  
Qy 3421 TTGCTGTGTTTGTGCTTTATTTGGTATGTTCCATCAGTATTTCTGTTTCACTTATATTTG 3480  
Db 3421 TTGCTGTGTTTGTGCTTTATTTGGTATGTTCCATCAGTATTTCTGTTTCACTTATATTTG 3480  
Qy 3481 CTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATCTG 3540  
Db 3481 CTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATCTATCTG 3540  
Qy 3541 TGGCAGCTTTGCTTGTATTTGCAATCCTGAAATTAACCTTTCTTTATGGGATACACAATTTG 3600  
Db 3541 TGGCAGCTTTGCTTGTATTTGCAATCCTGAAATTAACCTTTCTTTATGGGATACACAATTTG 3600  
Qy 3601 CAACTATTTCTCATATGCTTTTGTATCATCATTTCCATCTATCCACTTTAGGTTGCT 3660  
Db 3601 CAACTATTTCTCATATGCTTTTGTATCATCATTTCCATCTATCCACTTTAGGTTGCT 3660  
Qy 3661 TGAATTTCTTTCTAATAAGATTTCTTGAAGAAATGTCAGAAATTTTGGACACCTATAATTC 3720  
Db 3661 TGAATTTCTTTCTAATAAGATTTCTTGAAGAAATGTCAGAAATTTTGGACACCTATAATTC 3720  
Qy 3721 CATGGATAGGCTTTTCACTAGCTGTTTATATCGCTTTTACTGCAAGTGTGTACTGTGATTT 3780  
Db 3721 CATGGATAGGCTTTTCACTAGCTGTTTATATCGCTTTTACTGCAAGTGTGTACTGTGATTT 3780  
Qy 3781 TCCCTCTTCAATPACTATGAGAAATATATGGAGGAGATCAATTAAGAAAGATCCCTTTT 3840  
Db 3781 TCCCTCTTCAATPACTATGAGAAATATATGGAGGAGATCAATTAAGAAAGATCCCTTTT 3840  
Qy 3841 TCAGAAACCTTTCAACGAAAGTCTTAAATATAGGAAGCTTCCAGAACCCACAGCAATGAGG 3900  
Db 3841 TCAGAAACCTTTCAACGAAAGTCTTAAATATAGGAAGCTTCCAGAACCCACAGCAATGAGG 3900  
Qy 3901 ATGAAGATGAAGTGTCAAGCTGAAGACTTAAAGTCAAGAGCTGATGGTGGCTGCACT 3960  
Db 3901 ATGAAGATGAAGTGTCAAGCTGAAGACTTAAAGTCAAGAGCTGATGGTGGCTGCACT 3960  
Qy 3961 GTTGTGAGGAGAAACCATCCATTTATGCTCAGCAATTTTGCATTAAGAAATATGATGACAAG 4020  
Db 3961 GTTGTGAGGAGAAACCATCCATTTATGCTCAGCAATTTTGCATTAAGAAATATGATGACAAG 4020  
Qy 4021 AAGATTTCTTTCTTCAAGAAAGTGAAGAGTGGCACTAATACATCTCTTCTGTTG 4080  
Db 4021 AAGATTTCTTTCTTCAAGAAAGTGAAGAGTGGCACTAATACATCTCTTCTGTTG 4080  
Qy 4081 TGAAGAAAGGAGAGATCTTTAGGACTATTTGGTCCAAATGTTGCTGGCAAAAGCACAATTA 4140  
Db 4081 TGAAGAAAGGAGAGATCTTTAGGACTATTTGGTCCAAATGTTGCTGGCAAAAGCACAATTA 4140  
Qy 4141 TTAATATTTCTGTTGTTGATTTGAACCACTTTCAGGCCAGGATTTTTTTAGGAGATTTT 4200  
Db 4141 TTAATATTTCTGTTGTTGATTTGAACCACTTTCAGGCCAGGATTTTTTTAGGAGATTTT 4200



QY 517 TGCCCTATGAACCTCGTTTTTCTTGATATGATCCAGTATCTTCTATTTATATGGATT 576  
Db 1414 TGCCCTATGAACCTCGTTTTTCTTGATATGATCCAGTATCTTCTATTTATATGGATT 1473  
QY 577 CAAGAGCTGGCTTCCAAATCATGTGAGGCTGCTCAGTACTGGTCTCAGTTTCACAG 636  
Db 1474 CAAGAGCTGGCTTCCAAATCATGTGAGGCTGCTCAGTACTGGTCTCAGTTTCACAG 1533  
QY 637 TTTTACAAGCATCCATAGATGCTGCCATTTATACAGTTGAAGACCAATGTTCTCTTTGGA 696  
Db 1534 TTTTACAAGCATCCATAGATGCTGCCATTTATACAGTTGAAGACCAATGTTCTCTTTGGA 1593  
QY 697 AGGAGCTGAGTCAACTAAAGCTGTTTATATGGAGAACTGCTGTTGAGAAATAGATA 756  
Db 1594 AGGAGCTGAGTCAACTAAAGCTGTTTATATGGAGAACTGCTGTTGAGAAATAGATA 1653  
QY 757 CTTTCCCCGAGGAGTAATTTTAAATATACCTAGTTTATAGCATTTTCACTTTTGGATACT 816  
Db 1654 CTTTCCCCGAGGAGTAATTTTAAATATACCTAGTTTATAGCATTTTCACTTTTGGATACT 1713  
QY 817 TTTTGGCAATTCATATCGTAGCAGAAAGAAAGAAAGAAATATAAGATATA 876  
Db 1714 TTTTGGCAATTCATATCGTAGCAGAAAGAAAGAAAGAAATATAAGATATA 1773  
QY 877 TGGGCTTCATGATCTGCTTTTGGCTTCTGCTTCTGCTGCTTCTATATACAGTTTAAATT 936  
Db 1774 TGGGCTTCATGATCTGCTTTTGGCTTCTGCTTCTGCTGCTTCTATATACAGTTTAAATT 1833  
QY 937 TTCTTATGTCCTTCTTATGGAGTCATTTGCACAGCTTCTTTGTTATTTCTCTCAAGTA 996  
Db 1834 TTCTTATGTCCTTCTTATGGAGTCATTTGCACAGCTTCTTTGTTATTTCTCTCAAGTA 1893  
QY 997 GCAGCATTCGTATATTCGCTTTTCTTCTTATGGATATCATCTGATTTTGGCTT 1056  
Db 1894 GCAGCATTCGTATATTCGCTTTTCTTCTTATGGATATCATCTGATTTTGGCTT 1953  
QY 1057 TAATGCTGACACCTCTTTTAAAAAATCAAAACATGTGGATAGTTGAAATTTTGTGA 1116  
Db 1954 TAATGCTGACACCTCTTTTAAAAAATCAAAACATGTGGATAGTTGAAATTTTGTGA 2013  
QY 1117 CTGTGGCTTTTGAATTTATGGCTTTATGATTAATCTCATGAAAGTTTTCGCAATCGT 1176  
Db 2014 CTGTGGCTTTTGAATTTATGGCTTTATGATTAATCTCATGAAAGTTTTCGCAATCGT 2073  
QY 1177 TAGTGGCTTTTCACTGCTTCTGTCACGTGACCTTTTGTGATGCTATTCACAGTCA 1236  
Db 2074 TAGTGGCTTTTCACTGCTTCTGTCACGTGACCTTTTGTGATGCTATTCACAGTCA 2133  
QY 1237 TGCAATTAGAAGATTTAATGAAGTGTCTTCAATTTTCAAAATTTGACTCGAGGCCATATC 1296  
Db 2134 TGCAATTAGAAGATTTAATGAAGTGTCTTCAATTTTCAAAATTTGACTCGAGGCCATATC 2193  
QY 1297 CTCTAATTAATCAATATCATGCTCACAATTAATAGTATATCTCATGCTCTTGGCTG 1356  
Db 2194 CTCTAATTAATCAATATCATGCTCACAATTAATAGTATATCTCATGCTCTTGGCTG 2253  
QY 1357 TCTATCTGATCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 1416  
Db 2254 TCTATCTGATCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 2313  
QY 1417 TGAAGCTTCATATTTGCTCAAGAGCAAAAGAAATTTAGGAGTTATCAGAGGCCATG 1476  
Db 2314 TGAAGCTTCATATTTGCTCAAGAGCAAAAGAAATTTAGGAGTTATCAGAGGCCATG 2373  
QY 1477 TTAATGGAAATATAGTTTATAGTAAATTTAGTGGCTTCTTCAAGATTTGTAGGAA 1536  
Db 2374 TTAATGGAAATATAGTTTATAGTAAATTTAGTGGCTTCTTCAAGATTTGTAGGAA 2433  
QY 1537 AAGAAGCCATAGAATTAGTGATTTCAAGAGCATACAGAAAGAGGCTGAAATGTGG 1596  
Db 2434 AAGAAGCCATAGAATTAGTGATTTCAAGAGCATACAGAAAGAGGCTGAAATGTGG 2493

QY 1597 AGGCTTTGAGAAATTTGTCTATTTGACATATATGAGGCTCAGATTACTGCTTACTTGGCC 1656  
Db 2494 AGGCTTTGAGAAATTTGTCTATTTGACATATATGAGGCTCAGATTACTGCTTACTTGGCC 2553  
QY 1657 ACAGTGGAAACAGAAAGAGTACATTTGATGAATTTCTTTGTGACCTGCGCCACTTCTG 1716  
Db 2554 ACAGTGGAAACAGAAAGAGTACATTTGATGAATTTCTTTGTGACCTGCGCCACTTCTG 2613  
QY 1717 ATGGGTTTGATCTATATATGAGCAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAA 1776  
Db 2614 ATGGGTTTGATCTATATATGAGCAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAA 2673  
QY 1777 GAAAAATGATGTCATTTGTCACAGTATAGTATACACTTTGATGTTTGGACAGTAGAAG 1836  
Db 2674 GAAAAATGATGTCATTTGTCACAGTATAGTATACACTTTGATGTTTGGACAGTAGAAG 2733  
QY 1837 AAAATTTATCAATTTTGGCTTCAATCAAAAGGGATAGCAGCCAAATATAATAACAAGAG 1896  
Db 2734 AAAATTTATCAATTTTGGCTTCAATCAAAAGGGATAGCAGCCAAATATAATAACAAGAG 2793  
QY 1897 TGCAAGAGGTTTCTATAGATTTAGACATGACAGACTATCAAGATATAACCAAGCTAAAAAT 1956  
Db 2794 TGCAAGAGGTTTCTATAGATTTAGACATGACAGACTATCAAGATATAACCAAGCTAAAAAT 2853  
QY 1957 TAAAGTGTGCTCAAAAAGAAAGCTGTCTATTAGAAATGCTGTTCTTGGGAAACCAAGA 2016  
Db 2854 TAAAGTGTGCTCAAAAAGAAAGCTGTCTATTAGAAATGCTGTTCTTGGGAAACCAAGA 2913  
QY 2017 TACTGCTGTAGATGAACCAACAGCTGGAAATGGAACCCCTGTTCTCGACATATTTGTATGA 2076  
Db 2914 TACTGCTGTAGATGAACCAACAGCTGGAAATGGAACCCCTGTTCTCGACATATTTGTATGA 2973  
QY 2077 ATCTTTTAAATACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGATG 2136  
Db 2974 ATCTTTTAAATACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGATG 3033  
QY 2137 AAGCTGACATTTCTTGACAGTAGGAAAGCTGTGTATATCACAGGAATGCTGAAATGTTG 2196  
Db 3034 AAGCTGACATTTCTTGACAGTAGGAAAGCTGTGTATATCACAGGAATGCTGAAATGTTG 3093  
QY 2197 GTTCTTCAATGTTCTTCAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
Db 3094 GTTCTTCAATGTTCTTCAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153  
QY 2257 ACAAATATTGTGACACAGAAATCTTTCTTCTACTGCTTAAACAAATATATACCTGGAGCTA 2316  
Db 3154 ACAAATATTGTGACACAGAAATCTTTCTTCTACTGCTTAAACAAATATATACCTGGAGCTA 3213  
QY 2317 CTTTATTAACAAGATGACCAACAACTTGTGTATAGCTTGCCTTTCAAGGACATGGACA 2376  
Db 3214 CTTTATTAACAAGATGACCAACAACTTGTGTATAGCTTGCCTTTCAAGGACATGGACA 3273  
QY 2377 AATTTTCAGGTTTGTGTTCTGCTTACAGTATCATTTCAAAATTTGGGTGCTTCTTATG 2436  
Db 3274 AATTTTCAGGTTTGTGTTCTGCTTACAGTATCATTTCAAAATTTGGGTGCTTCTTATG 3333  
QY 2437 GTGTTTCCATGACGATTTTGGAAAGAGTATTTTAAAGCTAGAAAGTTGAAGCAAAATTG 2496  
Db 3334 GTGTTTCCATGACGATTTTGGAAAGAGTATTTTAAAGCTAGAAAGTTGAAGCAAAATTG 3393  
QY 2497 ACCAAGCAGATTTAGTGTATTTACTCAGCAGCCACTGGAGGAAGAAATGATTTCAAAAT 2556  
Db 3394 ACCAAGCAGATTTAGTGTATTTACTCAGCAGCCACTGGAGGAAGAAATGATTTCAAAAT 3453  
QY 2557 CTTTGTGATGAATTTGAAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
Db 3454 CTTTGTGATGAATTTGAAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 3513  
QY 2617 GCACCATGAGCTTTTGGAAACACAGATGTATATACATAGCAAGTTTCAATTTCTTTACCT 2676  
Db 3514 GCACCATGAGCTTTTGGAAACACAGATGTATATACATAGCAAGTTTCAATTTCTTTACCT 3573  
QY 2677 TGAAGCTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736

3574 TGAAACGCGTAAAGTAAATCAGTGAAGTCTGCTTCTGCTTTTAAATTTTTCACAG 3633  
 2737 TTCCAGATTTTATGTTTGTGTTTCTATCACTCTTTTAAATGCTGTTTCCCATCAAC 2796  
 3634 TTCCAGATTTTATGTTTGTGTTTCTATCACTCTTTTAAATGCTGTTTCCCATCAAC 3693  
 2797 TTGTTCCAGACTTATATTTTCTAAACCTGGAGACAAACCATATAAATAAACAAGTC 2856  
 3694 TTGTTCCAGACTTATATTTTCTAAACCTGGAGACAAACCATATAAATAAACAAGTC 3753  
 2857 TGCTTTCTCAAAATCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAGCC 2916  
 3754 TGCTTTCTCAAAATCTGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAGCC 3813  
 2917 AGAACATAATGTTGAGATGATTAAATGACAGTACTATGATCCGCTGCTCCCATAGTG 2976  
 3814 AGAACATAATGTTGAGATGATTAAATGACAGTACTATGATCCGCTGCTCCCATAGTG 3873  
 2977 CGGCTTTAAATGTTGATGCAATCGAAAGGACTATGTTTTCGAGCTGTTTTCACAGTA 3036  
 3874 CGGCTTTAAATGTTGATGCAATCGAAAGGACTATGTTTTCGAGCTGTTTTCACAGTA 3933  
 3037 CTATGCTTTATCTTTACCTATATTAAGTGAATCATTAAGTAACTACTATCTTTATCAT 3096  
 3934 CTATGCTTTATCTTTACCTATATTAAGTGAATCATTAAGTAACTACTATCTTTATCAT 3993  
 3097 TAAATGTGCTGAAACCATCCAGATCTGGAGTACCCATCTCTTCAAGAAATTAATGATA 3156  
 3994 TAAATGTGCTGAAACCATCCAGATCTGGAGTACCCATCTCTTCAAGAAATTAATGATA 4053  
 3157 TAGTTTTTAAATGTTGAGCTGATTATTTCAAGAGCTTTGCTTGGAACTATTGTTACTGCAA 3216  
 4054 TAGTTTTTAAATGTTGAGCTGATTATTTCAAGAGCTTTGCTTGGAACTATTGTTACTGCAA 4113  
 3217 TGCCACCTTACTTTGCAATGGAATAATGACAGAAATCAAGATCAAAAGCTTATATCAAC 3276  
 4114 TGCCACCTTACTTTGCAATGGAATAATGACAGAAATCAAGATCAAAAGCTTATATCAAC 4173  
 3277 TTAACCTTTACAGCTCTTTTGCATCTGCAATTTGGATTTGGAACAGCTGTTGTTGATATCC 3336  
 4174 TTAACCTTTACAGCTCTTTTGCATCTGCAATTTGGATTTGGAACAGCTGTTGTTGATATCC 4233  
 3337 CCTTATTTTATCACTTATTTGATGCTAGGAGCTTATTTGCAATTTCAATTTAGAT 3396  
 4234 CCTTATTTTATCACTTATTTGATGCTAGGAGCTTATTTGCAATTTCAATTTAGAT 4293  
 3397 TATATTTTATCACTTAAAGTTCCTTCTGCTGCTGTTTTCCTTATTTGTTATGTTCCAT 3456  
 4294 TATATTTTATCACTTAAAGTTCCTTCTGCTGCTGTTTTCCTTATTTGTTATGTTCCAT 4353  
 3457 CAGTTATTTCTGTTCACTTATTTGCTTCTTTTCAACCTTTTAAAGAAATTTTAAATACCAAG 3516  
 4354 CAGTTATTTCTGTTCACTTATTTGCTTCTTTTCAACCTTTTAAAGAAATTTTAAATACCAAG 4413  
 3517 AATTTTGGTCACTTATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3576  
 4414 AATTTTGGTCACTTATCTATCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4473  
 3577 CTTTCTTTATGGGATACAAATTTGCAATTTCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCT 3636  
 4474 CTTTCTTTATGGGATACAAATTTGCAATTTCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCT 4533  
 3637 CAATCTATCCACTTTAGGTTGCTGATTTCTTTTCAAAAGATTTCTTGGAAAGATGTAC 3696  
 4534 CAATCTATCCACTTTAGGTTGCTGATTTCTTTTCAAAAGATTTCTTGGAAAGATGTAC 4593  
 3697 GAAATAATGTTGACACTTAAATCTCAATGGATAGGCTTTTCACTAGCTGTTTATCGCCTT 3756  
 4594 GAAATAATGTTGACACTTAAATCTCAATGGATAGGCTTTTCACTAGCTGTTTATCGCCTT 4653  
 3757 ACCTGAGTGTGACTGTGATTTTCTCTTCAATATCTATGAGAAATAATGAGGCA 3816

4654 ACCTGAGTGTGACTGTGGAATTTCTCTTACAATACTATGAGAAAAATATGAGGCA 4713  
 3817 GATCAATTAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTTAAAAATAGGAAGC 3876  
 4714 GATCAATTAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTTAAAAATAGGAAGC 4773  
 3877 TTCCAGAACCCAGCAATGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 3936  
 4774 TTCCAGAACCCAGCAATGAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAG 4833  
 3937 TCAGAGAGCTGATGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3996  
 4834 TCAGAGAGCTGATGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4893  
 3997 TGCAATAAAGATATGATGACAAAGAAATTTTCTTTTCAAGAAAGTAAAGAAAGTGG 4056  
 4894 TGCAATAAAGATATGATGACAAAGAAATTTTCTTTTCAAGAAAGTAAAGAAAGTGG 4953  
 4057 CAATTAATAATCATCTCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4116  
 4954 CAATTAATAATCATCTCTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5013  
 4117 ATGCT 4176  
 5014 ATGCT 5073  
 4177 GCCAGGTATTTTAGGAGATTTTCTTTCAGAGACAAAGTGAAGATGATGATGATGATGATGATGATGATGATGAT 4236  
 5074 GCCAGGTATTTTAGGAGATTTTCTTTCAGAGACAAAGTGAAGATGATGATGATGATGATGATGATGATGATGAT 5133  
 4237 GTATGGGTACTGCTCTCCTCAGATAAACCCTTTTGTGGCCAGATACATACATTTGCAGGAACAT 4296  
 5134 GTATGGGTACTGCTCTCAGATAAACCCTTTTGTGGCCAGATACATACATTTGCAGGAACAT 5193  
 4297 TTGAAATTTTATGAGAGCTGCTCAAGAGATGAGTGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 4356  
 5194 TTGAAATTTTATGAGAGCTGCTCAAGAGATGAGTGCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5253  
 4357 GAATAACACATGACACTTGTATTTTAAAGAAACATCTTCAAGAAAGTGAAGATGATGATGATGATGATGATGATGATGATGAT 4416  
 5254 GAATAACACATGACACTTGTATTTTAAAGAAACATCTTCAAGAAAGTGAAGATGATGATGATGATGATGATGATGATGATGAT 5313  
 4417 CAGGAATCAAAACGAAAGTGTGTTTGTCTTAAGTATGCTAGGAAATCTCAGATTAATCTT 4476  
 5314 CAGGAATCAAAACGAAAGTGTGTTTGTCTTAAGTATGCTAGGAAATCTCAGATTAATCTT 5373  
 4477 TGCTAGATGAACCATCTACAGGTATGATCCCAAGCCAAACAGCACATGCTGGCGAGCAA 4536  
 5374 TGCTAGATGAACCATCTACAGGTATGATCCCAAGCCAAACAGCACATGCTGGCGAGCAA 5433  
 4537 TTCCAGACTGCAATTTAAACAGAGAGGCTGCTATTTCTGACCACTACTATATGAGG 4596  
 5434 TTCCAGACTGCAATTTAAACAGAGAGGCTGCTATTTCTGACCACTACTATATGAGG 5493  
 4597 AGGCAGAGGCT 4656  
 5494 AGGCAGAGGCT 5553  
 4657 GAACAGTCAACATCTTAAAGAGTAAATTTTGGAAAGAGCTACTTTTGGAAATTTAAATTTGA 4716  
 5554 GAACAGTCAACATCTTAAAGAGTAAATTTTGGAAAGAGCTACTTTTGGAAATTTAAATTTGA 5613  
 4717 AGGACTGATGAGAAACCTTAGAGTACCGCTTCAAGAGAAATTCAGTATATTTTCC 4776  
 5614 AGGACTGATGAGAAACCTTAGAGTACCGCTTCAAGAGAAATTCAGTATATTTTCC 5673  
 4777 CAATGCAAGCCGTCAGGAAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTTAAGGAAG 4836  
 5674 CAATGCAAGCCGTCAGGAAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTTAAGGAAG 5733  
 4837 ATGTTTCACTGCTTTCACATCTTTTAAAGTGAAGAGCTAACATGCTTTTGGCCA 4896  
 5734 ATGTTTCACTGCTTTCACATCTTTTAAAGTGAAGAGCTAACATGCTTTTGGCCA 5793



QY 517 TGTCTATGAACCTTGGTTTTTCTGATGATCCAGTATCTCTATTTATATGATTT 576  
Db 1414 TGTCTATGAACCTTGGTTTTTCTGATGATCCAGTATCTCTATTTATATGATTT 1473  
QY 577 CAAGAGCTGGCTTTCAAAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTCACAG 636  
Db 1474 CAAGAGCTGGCTTTCAAAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTCACAG 1533  
QY 637 TTTTCAAGCATCCATAGATGCTGCCATATACAGTTCGAGGCAATGTTCTCTTTGGA 696  
Db 1534 TTTTCAAGCATCCATAGATGCTGCCATATACAGTTCGAGGCAATGTTCTCTTTGGA 1593  
QY 697 AGGAGCTGGAGTCAACTAAAGCTGTTTATATGGGAGAACTGCTGTTGTAGAAATAGATA 756  
Db 1594 AGGAGCTGGAGTCAACTAAAGCTGTTTATATGGGAGAACTGCTGTTGTAGAAATAGATA 1653  
QY 757 CTTTCCCGGAGGAGTAATTTAATATACCTAGTATTTATAGCATTTTCACTTTTGGATCT 816  
Db 1654 CTTTCCCGGAGGAGTAATTTAATATACCTAGTATTTATAGCATTTTCACTTTTGGATCT 1713  
QY 817 TTTTGGCAATTCATATCGTAGCAAAAGAAAGAAAGAAATTTTAAAGATAA 876  
Db 1714 TTTTGGCAATTCATATCGTAGCAAAAGAAAGAAAGAAATTTTAAAGATAA 1773  
QY 877 TGGAGCTTCATGATCTGCTTTTGGCTTTCTCGGCTTCTCTATATACAAAGTTTAAATTT 936  
Db 1774 TGGAGCTTCATGATCTGCTTTTGGCTTTCTCGGCTTCTCTATATACAAAGTTTAAATTT 1833  
QY 937 TTTCTATGTCCTTTCTTATGGCAGTCAATTCGACAGCTTCTTTGTTATTTCTCTCAAGTA 996  
Db 1834 TTTCTATGTCCTTTCTTATGGCAGTCAATTCGACAGCTTCTTTGTTATTTCTCTCAAGTA 1893  
QY 997 GCAGCATGATGATTTCTGCTTTTCTTTTCTTTTATGATTAATCTGATTTCTTTGCTT 1056  
Db 1894 GCAGCATGATGATTTCTGCTTTTCTTTTCTTTTATGATTAATCTGATTTCTTTGCTT 1953  
QY 1057 TAATGTCGACCTCTTTTAAATCAAAACATGCGGATAGTCTGAATTTTGTGTA 1116  
Db 1954 TAATGTCGACCTCTTTTAAATCAAAACATGCGGATAGTCTGAATTTTGTGTA 2013  
QY 1117 CTGTGCTTTTGGATTTATGGCTTTATGATTAATCTCTATGAAAGTTTCCAAATCGT 1176  
Db 2014 CTGTGCTTTTGGATTTATGGCTTTATGATTAATCTCTATGAAAGTTTCCAAATCGT 2073  
QY 1177 TAGTGGCTTTTCACTGCTTTCTGCTCAGTACTTTTGTGATTTGGTATTCGACAGTCA 1236  
Db 2074 TAGTGGCTTTTCACTGCTTTCTGCTCAGTACTTTTGTGATTTGGTATTCGACAGTCA 2133  
QY 1237 TGCATTTAGAGATTTTAAATGAAGTGTCTTCAATTTTCAATTTGACTGCAGGCCCATATC 1296  
Db 2134 TGCATTTAGAGATTTTAAATGAAGTGTCTTCAATTTTCAATTTGACTGCAGGCCCATATC 2193  
QY 1297 CTCCTAATTTACATATGCTCACAACCTTAATAGTATATCTATGTCCTCTTGGCTG 1356  
Db 2194 CTCCTAATTTACATATGCTCACAACCTTAATAGTATATCTATGTCCTCTTGGCTG 2253  
QY 1357 TCTATCTTGATCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 1416  
Db 2254 TCTATCTTGATCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 2313  
QY 1417 TGAAGCTTCATATTTGGTCAAGAGCAAAAGAAATTTATGAGAGTTATCAGAGGCAATG 1476  
Db 2314 TGAAGCTTCATATTTGGTCAAGAGTAAAGAAATTTATGAGAGTTATCAGAGGCAATG 2373  
QY 1477 TTAATGGAATATTTAGTTTGTAGTAAATTTATGAGCCAGTTTCTTCAGAAATTTGAGGAA 1536  
Db 2374 TTAATGGAATATTTAGTTTGTAGTAAATTTATGAGCCAGTTTCTTCAGAAATTTGAGGAA 2433  
QY 1537 AAGAAGCCATPAGAATTTAGTGGTATTTCAAGAAGCATPAGAAGAGGGGTGAAATTTGCG 1596  
Db 2434 AAGAAGCCATPAGAATTTAGTGGTATTTCAAGAAGCATPAGAAGAGGGGTGAAATTTGCG 2493  
QY 1597 AGGCTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCC 1656

Db 2494 AGGCTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCC 2553  
QY 1657 ACAGTGGAAACAGAAAGAGTACATTTGATGAAATTTCTTTGTGGACTCTGCCCCACTTCTG 1716  
Db 2554 ACAGTGGAAACAGAAAGAGTACATTTGATGAAATTTCTTTGTGGACTCTGCCCCACTTCTG 2613  
QY 1717 ATGGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAA 1776  
Db 2614 ATGGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAA 2673  
QY 1777 GAAAAATGATTTGGCATTTTCTCCACAGTATATACACTTTGATGTTTGAAGTAGAAG 1836  
Db 2674 GAAAAATGATTTGGCATTTTCTCCACAGTATATACACTTTGATGTTTGAAGTAGAAG 2733  
QY 1837 AAAAAATGATTTGGCATTTTCTCCACAGTATATACACTTTGATGTTTGAAGTAGAAG 1896  
Db 2734 AAAAAATGATTTGGCATTTTCTCCACAGTATATACACTTTGATGTTTGAAGTAGAAG 2793  
QY 1897 TGCAAGAGGTTTTTCTAGATTTTAGACATGACATATCAAGATACCAAGCTTAAATAAT 1956  
Db 2794 TGCAAGAGGTTTTTCTAGATTTTAGACATGACATATCAAGATACCAAGCTTAAATAAT 2853  
QY 1957 TAAGTGGTGTCAAAAGAGAGTCTCTATTAGGAAATTCGTTCTTGGGAACCCAAAGA 2016  
Db 2854 TAAGTGGTGTCAAAAGAGAGTCTCTATTAGGAAATTCGTTCTTGGGAACCCAAAGA 2913  
QY 2017 TACTGCTGTAGTAAACCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGA 2076  
Db 2914 TACTGCTGTAGTAAACCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGA 2973  
QY 2077 ATCTTTTAAATPACAGAAAGCCAAATCGGGTGCAGTGTTCAGTACTCTTTCATGATG 2136  
Db 2974 ATCTTTTAAATPACAGAAAGCCAAATCGGGTGCAGTGTTCAGTACTCTTTCATGATG 3033  
QY 2137 AAGCTGACATCTTTCAGATAGGAAAGCTGTGATATCAAGGAATGCTGAAATGTTGTTG 2196  
Db 3034 AAGCTGACATCTTTCAGATAGGAAAGCTGTGATATCAAGGAATGCTGAAATGTTGTTG 3093  
QY 2197 GTTCTTCAATGTTCTTCAAAAGTAAATCGGGTGCAGTGTTCAGTACTCTTTCATGATG 2256  
Db 3094 GTTCTTCAATGTTCTTCAAAAGTAAATCGGGTGCAGTGTTCAGTACTCTTTCATGATG 3153  
QY 2257 ACAAATATTTGCGCACAGATCTTCTTCTCACTGGTTAAACAAACATATACCTGGAGCTA 2316  
Db 3154 ACAAATATTTGCGCACAGATCTTCTTCTCACTGGTTAAACAAACATATACCTGGAGCTA 3213  
QY 2317 CTTTATTAACAAGATGACCAACAACTTGTGTATAGTGTGCTTTCAAGGACATGACAA 2376  
Db 3214 CTTTATTAACAAGATGACCAACAACTTGTGTATAGTGTGCTTTCAAGGACATGACAA 3273  
QY 2377 AATTTTCAGGTTTGTCTTCTGCGCTAGACAGTCAATTTCAAAATTTGGGTGTCTTCTATG 2436  
Db 3274 AATTTTCAGGTTTGTCTTCTGCGCTAGACAGTCAATTTCAAAATTTGGGTGTCTTCTATG 3333  
QY 2437 GTGTTTCCATGACGACTTTTGGAAAGAGCTATTTTTAAAGCTTAGAAGTTGAAGCAAAATG 2496  
Db 3334 GTGTTTCCATGACGACTTTTGGAAAGAGCTATTTTTAAAGCTTAGAAGTTGAAGCAAAATG 3393  
QY 2497 ACCAAGCAGATTTAGTGTATTTTACTCAGCAGCCCTGAGGAGGAAATCGAATTTCAAAAT 2556  
Db 3394 ACCAAGCAGATTTAGTGTATTTTACTCAGCAGCCCTGAGGAGGAAATCGAATTTCAAAAT 3453  
QY 2557 CTTTGTGATGAAATGGAAACAGAGCTTACTTTATTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
Db 3454 CTTTGTGATGAAATGGAAACAGAGCTTACTTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 3513  
QY 2617 GCACCATGAGCCTTTGAAACCAACAGATGTATACATAGCAAAAGTTTCATTTCTTTACCT 2676  
Db 3514 GCACCATGAGCCTTTGAAACCAACAGATGTATACATAGCAAAAGTTTCATTTCTTTACCT 3573  
QY 2677 TGAACCTGTAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG 2736



Db 3574 TGAACGCTGAAAGTAAATCAGTGAGATCAGTGTCTGCTTTTAATTTTTTTCACAG 3633  
Qy 2737 TTCCAGATTTTATGTTTTTGGTTTATCACTCTTTTAAATGCTGTGTTCCATCAAC 2796  
Db 3634 TTCCAGATTTTATGTTTTTGGTTTATCACTCTTTTAAATGCTGTGTTCCATCAAC 3693  
Qy 2797 TTGTTCCAGACTTATATTTTCTTAAACCTGGAGACAAACCAATATAACAAACAAAGTC 2856  
Db 3694 TTGTTCCAGACTTATATTTTCTTAAACCTGGAGACAAACCAATATAACAAACAAAGTC 3753  
Qy 2857 TGCCTTTTCAAAATTTCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 2916  
Db 3754 TGCCTTTTCAAAATTTCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 3813  
Qy 2917 AGAACATAATGCTGACGATGATTATATGACAGTACTATGATCCGTGCTCCCATAGTG 2976  
Db 3814 AGAACATAATGCTGACGATGATTATATGACAGTACTATGATCCGTGCTCCCATAGTG 3873  
Qy 2977 CGGCTTTAAATGATGATGATTCAGAAAAGGACTATGTTTTTGGAGCTGTTTTCAACAGTA 3036  
Db 3874 CGGCTTTAAATGATGATGATTCAGAAAAGGACTATGTTTTTGGAGCTGTTTTCAACAGTA 3093  
Qy 3037 CTATGGTTTATTTCTTTACTATATTAAGTGAATATCATTAAGTAACTACTATCTTTATCAT 3096  
Db 3934 CTATGGTTTATTTCTTTACTATATTAAGTGAATATCATTAAGTAACTACTATCTTTATCAT 3093  
Qy 3097 TAAATGTAAGTGAACCACTCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAAGTGA 3156  
Db 3994 TAAATGTAAGTGAACCACTCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAAGTGA 4053  
Qy 3157 TAGTTTTTAAATTTAGCTGTATTTTCAAGAGCTTTGCTTGGATCATTTGTTACTGCA 3216  
Db 4054 TAGTTTTTAAATTTAGCTGTATTTTCAAGAGCTTTGCTTGGATCATTTGTTACTGCA 4113  
Qy 3217 TGCCACCTTACTTTGCCATGAAATGCAGAGAAATCAAGATCAAGCTTATTAAGTCAAC 3276  
Db 4114 TGCCACCTTACTTTGCCATGAAATGCAGAGAAATCAAGATCAAGCTTATTAAGTCAAC 4173  
Qy 3277 TTAACCTTTAGCTTTTGGCATCTGCATATGGAATGGAAGCTGTTGTTGATATCC 3336  
Db 4174 TTAACCTTTAGCTTTTGGCATCTGCATATGGAATGGAAGCTGTTGTTGATATCC 4233  
Qy 3337 CCTTATTTTATCACTTATTTTATGCTAGGAAGCTTATGSCATTTTCATTTATGGAT 3396  
Db 4234 CCTTATTTTATCACTTATTTTATGCTAGGAAGCTTATGSCATTTTCATTTATGGAT 4293  
Qy 3397 TATATTTTATCTATAAGTTCTTGTGCTGTTTGTGCTTTTGTGCTTATGTTATGTTCCAT 3456  
Db 4294 TATATTTTATCTATAAGTTCTTGTGCTGTTTGTGCTTTTGTGCTTATGTTATGTTCCAT 4353  
Qy 3457 CAGTTATTTCTGTTCACTTATATGCTTCTTCACTTAAAGAAATTTTAAATACCAAG 3516  
Db 4354 CAGTTATTTCTGTTCACTTATATGCTTCTTCACTTAAAGAAATTTTAAATACCAAG 4413  
Qy 3517 AATTTTGTGCTATTTATCTATTTCTGTCGAGCGTTGCTGTTGTTGCTTATGCAATCACTGAAATAA 3576  
Db 4414 AATTTTGTGCTATTTATCTATTTCTGTCGAGCGTTGCTGTTGCTTATGCAATCACTGAAATAA 4473  
Qy 3577 CTTTCTTTATGGATACAAATGCAATTCATATTCATATGCTTTTGTATCATCTTC 3636  
Db 4474 CTTTCTTTATGGATACAAATGCAATTCATATTCATATGCTTTTGTATCATCTTC 4533  
Qy 3637 CAATCTATCCACTTCTAGTGTGCTGATTTCTTTTCAATAAGATTTCTTGGAAAGTATGAC 3696  
Db 4534 CAATCTATCCACTTCTAGTGTGCTGATTTCTTTTCAATAAGATTTCTTGGAAAGTATGAC 4593  
Qy 3697 GAAATATGTTGACACCTATATATCAATGGATAGCTTTTCAAGTGTGTTATATCGGCTT 3756  
Db 4594 GAAATATGTTGACACCTATATATCAATGGATAGCTTTTCAAGTGTGTTATATCGGCTT 4653  
Qy 3757 ACCTGAGTGTACTGTGGATTTTCTTCTTACAAATATCTATGAGAAATATGAGGCA 3816  
Db 4654 ACCTGAGTGTACTGTGGATTTTCTTCTTACAAATATCTATGAGAAATATGAGGCA 4713

Qy 3817 GATCAATTAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAGTCTTAAATATGGAAC 3876  
Db 4714 GATCAATTAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAGTCTTAAATATGGAAC 4773  
Qy 3877 TTCCAGAACCCACAGCAATGAGGATGAAGATGCAAGATGTCAAAGCTGAAAGACTAAAGG 3936  
Db 4774 TTCCAGAACCCACAGCAATGAGGATGAAGATGCAAGATGTCAAAGCTGAAAGACTAAAGG 4833  
Qy 3937 TCAAGAGCTGATGGGTGCCAGTGTGTGAGGAGAAACCATCCATATATGTCAGCAAT 3996  
Db 4834 TCAAGAGCTGATGGGTGCCAGTGTGTGAGGAGAAACCATCCATATATGTCAGCAAT 4893  
Qy 3997 TGCATATAAGATATGATCAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGG 4056  
Db 4894 TGCATATAAGATATGATCAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGG 4953  
Qy 4057 CAACTAAATACATCTCTTTCTGTGTGAAAGAAAGGAGAGATCTTTAGGACTATTTGGTCCAA 4116  
Db 4954 CAACTAAATACATCTCTTTCTGTGTGAAAGAAAGGAGAGATCTTTAGGACTATTTGGTCCAA 5013  
Qy 4117 ATGGTCTGCGAAAGACAAATTTATTAATTTCTGTGTGCTGATATTTGAACCAACTTCAG 4176  
Db 5014 ATGGTCTGCGAAAGACAAATTTATTAATTTCTGTGTGCTGATATTTGAACCAACTTCAG 5073  
Qy 4177 GCCAGGATTTTTTGGAGATTTATTTCTCAGAGACAAAGTGAAGATGATTTCACTGAAGT 4236  
Db 5074 GCCAGGATTTTTTGGAGATTTATTTCTCAGAGACAAAGTGAAGATGATTTCACTGAAGT 5133  
Qy 4237 GATGGGTTTACTGCTCCCTCAGATATAACCTTTTGGCCAGATACTCATTTGCAAGCAAT 4296  
Db 5134 GATGGGTTTACTGCTCCCTCAGATATAACCTTTTGGCCAGATACTCATTTGCAAGCAAT 5193  
Qy 4297 TTGAAATTTTGGAGCTGTCAAAGGAATGAGTGCAGTGCATGAAGAAAGTCAATAGTC 4356  
Db 5194 TTGAAATTTTGGAGCTGTCAAAGGAATGAGTGCAGTGCATGAAGAAAGTCAATAGTC 5253  
Qy 4357 GAATAACATGCACTTTGATTTTAAAGAACTTTTCAAGAGACTGTTAAAGAACTTACCTG 4416  
Db 5254 GAATAACATGCACTTTGATTTTAAAGAACTTTTCAAGAGACTGTTAAAGAACTTACCTG 5313  
Qy 4417 CAGGAATCAAAAGAAAGTTGTGTTTTGCTCTAAGTATGCTAGGGAATCTCTCAGATTTACT 4476  
Db 5314 CAGGAATCAAAAGAAAGTTGTGTTTTGCTCTAAGTATGCTAGGGAATCTCTCAGATTTACT 5373  
Qy 4477 TGCTAGATGAACCATCTCAGGTAATGATCCCAAGCCAAACAGCACTGTGGCGAGCA 4536  
Db 5374 TGCTAGATGAACCATCTCAGGTAATGATCCCAAGCCAAACAGCACTGTGGCGAGCA 5433  
Qy 4537 TTGCAACTGCTTTTAAAGAAAGCGGCTGCTTATTTCTGACCACTCACTATATGGAGG 4596  
Db 5434 TTGCAACTGCTTTTAAAGAAAGCGGCTGCTTATTTCTGACCACTCACTATATGGAGG 5493  
Qy 4597 AGCGAGAGCTGTCTGTGATTCGAGTAGCTATCATGTGTCTGGCGAGTTAAGATGATCG 4656  
Db 5494 AGCGAGAGCTGTCTGTGATTCGAGTAGCTATCATGTGTCTGGCGAGTTAAGATGATCG 5553  
Qy 4657 GAACAGTCAACATCTTAAAGAGTAAATTTGGAAGAGGCTTCTTTTGGAAATTTAAATTCGA 4716  
Db 5554 GAACAGTCAACATCTTAAAGAGTAAATTTGGAAGAGGCTTCTTTTGGAAATTTAAATTCGA 5613  
Qy 4717 AGGACTGGATAGAAACCTTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCC 4776  
Db 5614 AGGACTGGATAGAAACCTTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCC 5673  
Qy 4777 CAAATCGAGCGTCAGGAAAGTTTTTCTTATTTTGGCTTATATAAATTCCTAAGGAAG 4836  
Db 5674 CAAATCGAGCGTCAGGAAAGTTTTTCTTATTTTGGCTTATATAAATTCCTAAGGAAG 5733  
Qy 4837 ATGTTTCAGTCCCTTTTCACAATCTTTTTTAAAGTGGAGAAAGCTAAACATGCTTTTGCA 4896  
Db 5734 ATGTTTCAGTCCCTTTTCACAATCTTTTTTAAAGTGGAGAAAGCTAAACATGCTTTTGCA 5793



QY 4897 TTGAAGATATAGCTTTCTCAAGCAACATGGAAAGAGGTTTTGTAGAACTCACTAAAG 4956  
 DB 5794 TTGAAGATATAGCTTTCTCAAGCAACATGGAAAGAGGTTTTGTAGAACTCACTAAAG 5853  
 QY 4957 AACAGAGGAGGAGATATAGTTGTGGAACTTTAAACAGCAGCAGTTTGGTGGGACGAA 5016  
 DB 5854 AACAGAGGAGGAGATATAGTTGTGGAACTTTAAACAGCAGCAGTTTGGTGGGACGAA 5913  
 QY 5017 CACAAGAGATAGAGTAGTATTTTGAATTTGTATTTTGGTCTGCTTACTCGGACTTCT 5076  
 DB 5914 CACAAGAGATAGAGTAGTATTTTGAATTTGTATTTTGGTCTGCTTACTCGGACTTCT 5973  
 QY 5077 TTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGGTTTTTATTTGGAATGTGTAAGTGA 5136  
 DB 5974 TTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGGTTTTTATTTGGAATGTGTAAGTGA 6033  
 QY 5137 GAACCAAGACGACCTGAAATTTTCTAAGTCTCTTAATTTGAATTTGCTGTTGTG 5196  
 DB 6034 GAACCAAGACGACCTGAAATTTTCTAAGTCTCTTAATTTGAATTTGCTGTTGTG 6093  
 QY 5197 TTTTGTCTTTCTTTAAATAAAGCTATGTATTAATTAAGTGAAGCTGCTGTTGTATGA 5256  
 DB 6094 TTTTGTCTTTCTTTAAATAAAGCTATGTATTAATTAAGTGAAGCTGCTGTTGTATGA 6153  
 QY 5257 AGTATATGAATATAGTTTGTATGTATCTTTTCAATTTCAATTTCAATTTCAATTTCA 5316  
 DB 6154 AGTATATGAATATAGTTTGTATGTATCTTTTCAATTTCAATTTCAATTTCAATTTCA 6213  
 QY 5317 GAATTTGTGATTTAAAGGAACTGTAATAGAAAGTTTATTTTAAAGTTTATTTTAAAGTT 5376  
 DB 6214 GAATTTGTGATTTAAAGGAACTGTAATAGAAAGTTTATTTTAAAGTTTATTTTAAAGTT 6273  
 QY 5377 TATGCCATCTTTTAAATAAGTACGTAATGTTTCCATCTTAAATAAATAAATAAATAAATAA 5436  
 DB 6274 TATGCCATCTTTTAAATAAGTACGTAATGTTTCCATCTTAAATAAATAAATAAATAAATAA 6333  
 QY 5437 CTATGATAGAAAGATACATTAAGCAATGTGAAGTT 5475  
 DB 6334 CTATGATAGAAAGATACATTAAGCAATGTGAAGTT 6372

RESULT 4  
 AX537470  
 LOCUS AX537470 5463 bp DNA linear PAT 23-NOV-2002  
 DEFINITION Sequence 1 from Patent WO02070690.  
 ACCESSION AX537470  
 VERSION AX537470.1 GI:25269277  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1  
 AUTHORS Chen,H., Kilinski,L. and le Bihan,S.  
 TITLE Abca5 transporter and uses thereof  
 JOURNAL Patent: WO 02070690-A 1 12-SEP-2002;  
 Active Pass Pharmaceuticals, Inc. (CA)  
 FEATURES  
 source  
 1..5463  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Query Match 97.8%; Score 5356.6; DB 6; Length 5463;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 5440; Conservative 0; Mismatches 4; Indels 50; Gaps 3;  
 QY 1 GCCTCGCGCCCTCGCACAGATCCNAGCTGGTCAACCGCAGCTGAGTCAACAGACTCGAGC 60  
 DB 1 GCCTCGCGCCCTCGCACAGATCCNAGCTGGTCAACCGCAGCTGAGTCAACAGACTCGAGC 60  
 QY 61 GCCTCCAGGCTGACAGCTCTCGGCTCGGCGCTGAGGTTTATTCAGAAAACATGTCCA 120

DB 61 GCCTCCAGGCTGACAGCTCTCGGCTCGGCGCTGAGGTTTATTCAGAAAACATGTCCA 120  
 QY 121 CTGCAATTTAGGAGTATAGGTTTGGAGACAGACCAACACTTCTACTCAAGAAATTAAT 180  
 DB 121 CTGCAATTTAGGAGTATAGGTTTGGAGACAGACCAACACTTCTACTCAAGAAATTAAT 180  
 QY 181 TAATTAATGACAGAACCAAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTATTTTTTT 240  
 DB 181 TAATTAATGACAGAACCAAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTATTTTTTT 240  
 QY 241 TATTTTGGTTAATTAATTTAGCATGATGCATCCAAATAGAAATATAGAAAGAGTGCTA 300  
 DB 241 TATTTTGGTTAATTAATTTAGCATGATGCATCCAAATAGAAATATAGAAAGAGTGCTA 300  
 QY 301 ATATAGAACTCAATCTTATGGACAGATTTACTCTTTCTAATCTAATCTTTGGATATATC 360  
 DB 301 ATATAGAACTCAATCTTATGGACAGATTTACTCTTTCTAATCTAATCTTTGGATATATC 360  
 QY 361 CAGTGACTAATATTTACAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTACCTGATG 420  
 DB 361 CAGTGACTAATATTTACAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTACCTGATG 420  
 QY 421 TCATAATTTACTGAAGATATACAAATGAAAAGAAATGTTTAAACATCCAGTCTCTCTAAGC 480  
 DB 421 TCATAATTTACTGAAGATATACAAATGAAAAGAAATGTTTAAACATCCAGTCTCTCTAAGC 480  
 QY 481 CGACCACTTTGTAGGTGGTTTTTCAAGACTCCATGCTCTATGACTTCTGTTTTTTC 540  
 DB 481 CGACCACTTTGTAGGTGGTTTTTCAAGACTCCATGCTCTATGACTTCTGTTTTTTC 540  
 QY 541 CTGATATGATTTCCAGTATCTTTATTTATATGGATTCAAGAGCTGGTGTCAAAATCAT 600  
 DB 541 CTGATATGATTTCCAGTATCTTTATTTATATGGATTCAAGAGCTGGTGTCAAAATCAT 600  
 QY 601 GTGAGGCTGCTCAGTACTGCTCTCAGGTTTTCAGGTTTTCAGGATCCATAGATGCTG 660  
 DB 601 GTGAGGCTGCTCAGTACTGCTCTCAGGTTTTCAGGTTTTCAGGATCCATAGATGCTG 660  
 QY 661 CCAATTATACAGTTGAAGCAATGTTTCTCTTTGGAAGGAGCTGGAGTCAACTAAAGCTG 720  
 DB 661 CCAATTATACAGTTGAAGCAATGTTTCTCTTTGGAAGGAGCTGGAGTCAACTAAAGCTG 720  
 QY 721 TTATATGGAGAAATCTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTAAATTTAA 780  
 DB 721 TTATATGGAGAAATCTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTAAATTTAA 780  
 QY 781 TATACCTAGTTATAGCAATTTTACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG 840  
 DB 781 TATACCTAGTTATAGCAATTTTACCTTTTGGATACCTTTTGGCAATTCATATCGTAGCAG 840  
 QY 841 AAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATCTGCCTTTT 900  
 DB 841 AAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATCTGCCTTTT 900  
 QY 901 GGCTTTCTCGGTTCTTCTATATACAGTTTAAATTTTCTATGTCCTCTTTATGGCAG 960  
 DB 901 GGCTTTCTCGGTTCTTCTATATACAGTTTAAATTTTCTATGTCCTCTTTATGGCAG 960  
 QY 961 TCATTTGACAGCTCTCTTTGTTATTTTCTCAAAAGTAGCAGATTTGATATTTCTGCTTT 1020  
 DB 961 TCATTTGACAGCTCTCTTTGTTATTTTCTCAAAAGTAGCAGATTTGATATTTCTGCTTT 1020  
 QY 1021 TTTTCTTTATGATTATCATCTGTTATTTTCTTTTAAATGCTGACACCTCTTTTAAAAA 1080  
 DB 1021 TTTTCTTTATGATTATCATCTGTTATTTTCTTTTAAATGCTGACACCTCTTTTAAAAA 1080  
 QY 1081 AATCAAAACATGTGGGAATAGTTGAATTTTGTATCTGCTGCTTTTGGATTTATTTGGCC 1140  
 DB 1081 AATCAAAACATGTGGGAATAGTTGAATTTTGTATCTGCTGCTTTTGGATTTATTTGGCC 1140  
 QY 1141 TTAATGATAATCTTCATAGAAAGTTTCCCAAACTGTTAGTGGCTTTTCAGTCCCTTCT 1200

1141 TTATGATAATCTCTCAGAGAGAGTTTCCCAATCGTTAGTGGCTTTTCAGTCTCTTCT 1200  
 1201 GTCACTGTACTTTTGTGATGTTGATGACAGGTCATGATTTAGAGATTTTAAATGAAG 1260  
 1201 GTCACTGTACTTTTGTGATGTTGATGACAGGTCATGATTTAGAGATTTTAAATGAAG 1260  
 1261 GTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCCTCTAATATTACAAATATCATGC 1320  
 1261 GTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCCTCTAATATTACAAATATCATGC 1320  
 1321 TCACACTTAATAGTATATCTATGCTCTTGGCTGCTATCTGATCAAGTCAATTCAG 1380  
 1321 TCACACTTAATAGTATATCTATGCTCTTGGCTGCTATCTGATCAAGTCAATTCAG 1380  
 1381 GGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATGGTCAAGA 1440  
 1381 GGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATGGTCAAGA 1440  
 1441 GCAAAAGAAATTTAGAGGATTTATCAGAGGGCAATGTTAATGGAAATATTAGTTAGTG 1500  
 1441 GCAAAAGAAATTTAGAGGATTTATCAGAGGGCAATGTTAATGGAAATATTAGTTAGTG 1500  
 1501 AAATATTAGCAGGCTTCTTCTAGAAATTTGAGGAAAGAGCCCAATAGAAATTTAGTGTA 1560  
 1501 AAATATTAGCAGGCTTCTTCTAGAAATTTGAGGAAAGAGCCCAATAGAAATTTAGTGTA 1560  
 1561 TTCAGAGGACATACAGAAAGAGGGTGAATTTGAGAGGCTTTGAGAAATTTGCAATTTG 1620  
 1561 TTCAGAGGACATACAGAAAGAGGGTGAATTTGAGAGGCTTTGAGAAATTTGCAATTTG 1620  
 1621 ACATATATCAGGCTCAGATTTCTGCTTACTTGGCCACAGTGGAAACAGGAAAGATACAT 1680  
 1621 ACATATATCAGGCTCAGATTTCTGCTTACTTGGCCACAGTGGAAACAGGAAAGATACAT 1680  
 1681 TGATGAATATTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATATATGAC 1740  
 1681 TGATGAATATTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATATATGAC 1740  
 1741 ACAGAGCTTCAGAAATAGATGAATTTTGAAGCAAGAAATGATGGCAATTTGTCCAC 1800  
 1741 ACAGAGCTTCAGAAATAGATGAATTTTGAAGCAAGAAATGATGGCAATTTGTCCAC 1800  
 1801 AGTTAGATATACCTTTGATGTTTTCACAGTAGAGAAATTTATCAATTTTGGCTTCAA 1860  
 1801 AGTTAGATATACCTTTGATGTTTTCACAGTAGAGAAATTTATCAATTTTGGCTTCAA 1860  
 1861 TCAAGGGATACAGGCAATATATATCAAGAGATGCGAGGTTTACTAGATTTAG 1920  
 1861 TCAAGGGATACAGGCAATATATATCAAGAGATGCGAGGTTTACTAGATTTAG 1920  
 1921 ACATGCAGACTATCAAGATACCAAGCTTAAATAATTAAGTGGTCAAAAAGAAAGC 1980  
 1921 ACATGCAGACTATCAAGATACCAAGCTTAAATAATTAAGTGGTCAAAAAGAAAGC 1980  
 1981 TGTCATTAGGAATGCTGTTTGGGACCCAAAGATCTGCTGTAGATGAACCAAG 2040  
 1981 TGTCATTAGGAATGCTGTTTGGGACCCAAAGATCTGCTGTAGATGAACCAAG 2040  
 2041 CTGGAAATGACCCCTGCTCGACATATTTGATGGAATCTTTTAAATACAGAAAGCCA 2100  
 2041 CTGGAAATGACCCCTGCTCGACATATTTGATGGAATCTTTTAAATACAGAAAGCCA 2100  
 2101 ATCGGGTGCAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTGCAGATGGA 2160  
 2101 ATCGGGTGCAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTGCAGATGGA 2160  
 2161 AGCTGTGATATCAAGGAATGCTGAAATGTTGTTCTTCAATGTTCTCTCAAAAGTA 2220  
 2161 AGCTGTGATATCAAGGAATGCTGAAATGTTGTTCTTCAATGTTCTCTCAAAAGTA 2220  
 2221 AATGGGGATCGGCTACCGCTGAGCATGTAATAGACAAATATTTGTGCCACAGAAATCTC 2280  
 2221 AATGGGGATCGGCTACCGCTGAGCATGTAATAGACAAATATTTGTGCCACAGAAATCTC 2280

2281 TTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATTTACACAGAAATGACCAAC 2340  
 2281 TTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATTTACACAGAAATGACCAAC 2340  
 2341 AACTGTGTATAGCTTGCCTTTCAAGGACATGACAAATTTTTCAGGTTGTTTCTGCCC 2400  
 2341 AACTGTGTATAGCTTGCCTTTCAAGGACATGACAAATTTTTCAGGTTGTTTCTGCCC 2400  
 2401 TAGACAGTCAATCAAAATTTGGGTTGCTATTTCTTATGTTGTTTCCATGACGACTTTGGAAG 2460  
 2401 TAGACAGTCAATCAAAATTTGGGTTGCTATTTCTTATGTTGTTTCCATGACGACTTTGGAAG 2460  
 2461 ACCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTCACCAAGCAGATTTAGTGTATTTA 2520  
 2461 ACCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTCACCAAGCAGATTTAGTGTATTTA 2520  
 2521 CTGAGAGCCACCTGGAGGAAAGAAATCGATTCAAATCTTTTGTGAAATGGACAGAGCT 2580  
 2521 CTGAGAGCCACCTGGAGGAAAGAAATCGATTCAAATCTTTTGTGAAATGGACAGAGCT 2580  
 2581 TACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTTGAAACCAAC 2640  
 2581 TACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTTGAAACCAAC 2640  
 2641 AGATGTATACAAATAGCAAAAGTTTCAATTTCTTTACCTTGAAAGTGAATCAAGTGA 2700  
 2641 AGATGTATACAAATAGCAAAAGTTTCAATTTCTTTACCTTGAAAGTGAATCAAGTGA 2700  
 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTGTTT 2760  
 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTGTTT 2760  
 2761 ATCAGTCTTTTAAATAGTGTGTTTCCCAATCAAACTGTTTTCAGACTTATATTTCTAA 2820  
 2761 ATCAGTCTTTTAAATAGTGTGTTTCCCAATCAAACTGTTTTCAGACTTATATTTCTAA 2820  
 2821 AACCTGGAGCAAAACACATTAATAAACAAGCTGCTGCTTCTTCAAAATTTCTGCT --- 2876  
 2821 AACCTGGAGCAAAACACATTAATAAACAAGCTGCTGCTTCTTCAAAATTTCTGCT --- 2880  
 2877 -----GACTCAGATATCAGTGATCTTATAGCTTTTTCACAAGCCAGAAACA 2922  
 2881 AGAGTGTGTTGAAAGCTCAGATATCAGTGATCTTATAGCTTTTTCACAAGCCAGAAACA 2940  
 2923 TAATGTGAGCATGATTAATGACAGTATGATGATCCGTTGCTCCCATAGTGGGCTT 2982  
 2941 TAATGTGAGCATGATTAATGACAGTATGATGATCCGTTGCTCCCATAGTGGGCTT 3000  
 2983 TAAATGTGATGCTTTCAGAAAGGACTATGTTTTCAGCTGTTTTCACAAGCTATG 3042  
 3001 TAAATGTGATGCTTTCAGAAAGGACTATGTTTTCAGCTGTTTTCACAAGCTATG 3060  
 3043 TTTATTTCTTACCTATATTAGTGAATATCATTTAGTAACTACTTCTTTATCATTTAAATG 3102  
 3061 TTTATTTCTTACCTATATTAGTGAATATCATTTAGTAACTACTTCTTTATCATTTAAATG 3120  
 3103 TGACTGAAACCATCCAGATCTGGAGTACCCCATTTCTTTCAGAAATTTACTGATATAGTTT 3162  
 3121 TGACTGAAACCATCCAGATCTGGAGTACCCCATTTCTTTCAGAAATTTACTGATATAGTTT 3180  
 3163 TTAATAATTTAGCTGTATTTTCAAGAGCTTTGCTGG-AATCATTTGTTACTGCAATGCCA 3221  
 3181 TTAATAATTTAGCTGTATTTTCAAGAGCTTTGCTGGAAATCATTTAACTGCAATGCCA 3240  
 3222 CCTTACTTTGCCATGGAATGAGAGAAATCATTAAGATCAAAAGCTTTTACTCAACTTAA 3281  
 3241 CCTTACTTTGCCATGGAATGAGAGAAATCATTAAGATCAAAAGCTTTTACTCAACTTAA 3274  
 3282 CTTTCAGGCTTTTGGCATCTGCAATTTGGATTTGGAAGCTGTTGTTGATATCCCTTAA 3341  
 3275 -----AGGCTTTTGGCATCTGCAATTTGGATTTGGAAGCTGTTGTTGATATCCCTTAA 3329

3342 TTTTATCATCTCTATATTTTGATGCTAGGAAGCTTATTTGGCAATTCATATGGAATATAT 3401  
3330 TTTTATCATCTCTATATTTTGATGCTAGGAAGCTTATTTGGCAATTCATATGGAATATAT 3389  
3402 TTTTATCTGTAAGATTCCTGCTGGTATTTTGGCTTATTTGGTATATGTTTCCATCAGTT 3461  
3390 TTTTATCTGTAAGATTCCTGCTGGTATTTTGGCTTATTTGGTATATGTTTCCATCAGTT 3449  
3462 ATTCTGTTCACTATATTTGCTTTCTTACCTTTTAAAGAAAATTTTAAATACCAAGAAATTT 3521  
3450 ATTCTGTTCACTATATTTGCTTTCTTACCTTTTAAAGAAAATTTTAAATACCAAGAAATTT 3509  
3522 TGGTCATTTATCTATTTCTGTCGAGCGTTGGCTTGTATTGCAATCAGTGAATTAACCTTTC 3581  
3510 TGGTCATTTATCTATTTCTGTCGAGCGTTGGCTTGTATTGCAATCAGTGAATTAACCTTTC 3569  
3582 TTTATGGGATACACAATTCGAATTTCTTCATTTATTTGCTTTTGTATCATCATTTCCAAATC 3641  
3570 TTTATGGGATACACAATTCGAATTTCTTCATTTATTTGCTTTTGTATCATCATTTCCAAATC 3629  
3642 TATCCACTTCTAGTTTGGCTGATTTCTTTTCAATAAGATTTCTTGGGAAGATGTAGCAAAA 3701  
3630 TATCCACTTCTAGTTTGGCTGATTTCTTTTCAATAAGATTTCTTGGGAAGATGTAGCAAAA 3689  
3702 AATGTGGAACCTTATTAATCCATGGGATAGGCTTTTCTAGTAGCTGTTTATATCGCCTTACCTG 3761  
3690 AATGTGGAACCTTATTAATCCATGGGATAGGCTTTTCTAGTAGCTGTTTATATCGCCTTACCTG 3749  
3762 CAGTGTGATCTGGGATTTTCTCTTACAAATCTATGAGAAAATATGAGGAGGAGATCA 3821  
3750 CAGTGTGATCTGGGATTTTCTCTTACAAATCTATGAGAAAATATGAGGAGGAGATCA 3809  
3822 ATAAGAAAAGATCCCTTTTTCAGAAACCTTTTCAAGAAAGTCTTAAATAATGAGAACTTCCA 3881  
3810 ATAAGAAAAGATCCCTTTTTCAGAAACCTTTTCAAGAAAGTCTTAAATAATGAGAACTTCCA 3869  
3882 GAACCAACAGACATGAGATGAGATGAGATGCTTCAAGCTGGAAGCTTAAAGTCTAAA 3941  
3870 GAACCAACAGACATGAGATGAGATGAGATGCTTCAAGCTGGAAGCTTAAAGTCTAAA 3929  
3942 GAGCTGATGGTGTGCGAGTGTGTGAGGAGAAACCATCATTTATGGTCTAGCAATTTGCAAT 4001  
3930 GAGCTGATGGTGTGCGAGTGTGTGAGGAGAAACCATCATTTATGGTCTAGCAATTTGCAAT 3989  
4002 AAGAAATATGACAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAGTGGCAACT 4061  
3990 AAGAAATATGACAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAGTGGCAACT 4049  
4062 AATATACATCTTCTGCTGAAAAGAGAGAGATCTTAGGACTATTTGGTCTCAATGGT 4121  
4050 AATATACATCTTCTGCTGAAAAGAGAGAGATCTTAGGACTATTTGGTCTCAATGGT 4109  
4122 GCTGGCAAGACAAATTAATATCTGTTGGTGGTATTTGAAACCAATTTGAGGCCAG 4181  
4110 GCTGGCAAGACAAATTAATATCTGTTGGTGGTATTTGAAACCAATTTGAGGCCAG 4169  
4182 GTATTTTATAGGATTTATTTCTCAGAGACAGTGAAGATGATGATCTCAAGTGGTATG 4241  
4170 GTATTTTATAGGATTTATTTCTCAGAGACAGTGAAGATGATGATCTCAAGTGGTATG 4229  
4242 GGTACTGTCTCAGATAAACCTTTTGGCCAGATACCTACATTCAGGAAACATTTTCAA 4301  
4230 GGTACTGTCTCAGATAAACCTTTTGGCCAGATACCTACATTCAGGAAACATTTTCAA 4289  
4302 ATTTATGGAGCTGTCAAGAGATGATGCAAGTGCATGAAAGAGTCTAAGTGCATTA 4361  
4290 ATTTATGGAGCTGTCAAGAGATGATGCAAGTGCATGAAAGAGTCTAAGTGCATTA 4349  
4362 ACACATGCACTTGATTTTAAAGAAACATCTTTCAGAGACTGTAAAGAAACTTACCTCAGGA 4421  
4350 ACACATGCACTTGATTTTAAAGAAACATCTTTCAGAGACTGTAAAGAAACTTACCTCAGGA 4409  
4422 ATCAAGAAAGTGTGTTTGTCTTAAGTATGCTAGGGAATCTCAGATTTCTTGCTA 4481

RESULT 5

4410 ATCAAGCAAGATGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTTACTTTGCTA 4469  
4482 GATGAACCACTCTACAGTATGGATCCAAAGCCAAACAGCACATGTGGGAGCAATTCGA 4541  
4470 GATGAACCACTCTACAGTATGGATCCAAAGCCAAACAGCACATGTGGGAGCAATTCGA 4529  
4542 ACTGCAATTTAAACACAGAAAGCGGCTGTATTTCTGACCACTTCTATATGAGGAGGCA 4601  
4530 ACTGCAATTTAAACACAGAAAGCGGCTGTATTTCTGACCACTTCTATATGAGGAGGCA 4589  
4602 GAGGCTGTCTGTCATCGAGTAGCTATCATGGTGTCTGGGAGTAAAGATGTATCGGAACA 4661  
4590 GAGGCTGTCTGTCATCGAGTAGCTATCATGGTGTCTGGGAGTAAAGATGTATCGGAACA 4649  
4662 GTACAACTCTAAAGAGTAAATTTGAAAAGGCTACTTTTGGAAATTTAAATGAGGAC 4721  
4650 GTACAACTCTAAAGAGTAAATTTGAAAAGGCTACTTTTGGAAATTTAAATGAGGAC 4709  
4722 TGGATAGAAAACCTAGAGTAGACCCGCTTCAAGAGAAAATTCAGTATATTTTCCAAAT 4781  
4710 TGGATAGAAAACCTAGAGTAGACCCGCTTCAAGAGAAAATTCAGTATATTTTCCAAAT 4769  
4782 GCAAGCGCTCAGAAAAGTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATTT 4841  
4770 GCAAGCGCTCAGAAAAGTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATTT 4829  
4842 CAGTCCCTTTCACAACTCTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTGCATTTGAA 4901  
4830 CAGTCCCTTTCACAACTCTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTGCATTTGAA 4889  
4902 GAATATAGCTTTTCTCAAGCAATTTGGAACAGGTTTTTGTAGAACTCTAATAAGAACAA 4961  
4890 GAATATAGCTTTTCTCAAGCAATTTGGAACAGGTTTTTGTAGAACTCTAATAAGAACAA 4949  
4962 GAGGAGAGATTAATAGTTGTGGAACCTTTAAACAGCACACTTTTGGTGGAAACCAACAA 5021  
4950 GAGGAGAGATTAATAGTTGTGGAACCTTTAAACAGCACACTTTTGGTGGAAACCAACAA 5009  
5022 GAAGATAGAGTAGTATTTTGAATTTGTATTTGGTCTGCTTACTTGGGACTTCTTTCTT 5081  
5010 GAAGATAGAGTAGTATTTTGAATTTGTATTTGGTCTGCTTACTTGGGACTTCTTTCTT 5069  
5082 TTTTCACTTAATTTTAACTTTTGGTTTAAAGTTTTTTTTTATTTGGAAATGCTAAGTGAACC 5141  
5070 TTTTCACTTAATTTTAACTTTTGGTTTAAAGTTTTTTTTTATTTGGAAATGCTAAGTGAACC 5129  
5142 AAGAACGCACTTGAAATTTTCTTAAGCTCTTAAATGAAATGCTGTTGTTGTTTGTG 5201  
5130 AAGAACGCACTTGAAATTTTCTTAAGCTCTTAAATGAAATGCTGTTGTTGTTTGTG 5189  
5202 CTTTCTTTTAAATAAACGCTATGATATAATTAAGTGAAGTGCATGTTTGTATTGAAGTAT 5261  
5190 CTTTCTTTTAAATAAACGCTATGATATAATTAAGTGAAGTGCATGTTTGTATTGAAGTAT 5249  
5262 ATTGAACATAATATAGTTTGTATGTCACTTTTCCCAATTCAGAAAACAGTCTTCTGAAAT 5321  
5250 ATTGAACATAATATAGTTTGTATGTCACTTTTCCCAATTCAGAAAACAGTCTTCTGAAAT 5309  
5322 TGTGATTTAAAGAAATTTGAATAGATAGTATTTTATTTTAAAGTATCTTTTAAAGTTATGC 5381  
5310 TGTGATTTAAAGAAATTTGAATAGATAGTATTTTATTTTAAAGTATCTTTTAAAGTTATGC 5369  
5382 CATCTTCTTAAATAAGTACGTAATGTTTCCAATCTAAATAAAAAAATAATTCATACTAAT 5441  
5370 CATCTTCTTAAATAAGTACGTAATGTTTCCAATCTAAATAAAAAAATAATTCATACTAAT 5429  
5442 GCATAGAAAAGATACATAAAGCAATGTGAAGTT 5475  
5430 GCATAGAAAAGATACATAAAGCAATGTGAAGTT 5463

AX392931  
LOCUS AX392931 6369 bp DNA linear PAT 23-MAR-2002  
DEFINITION Sequence 33 from Patent WO0212340.  
ACCESSION AX392931  
VERSION AX392931.1 GI:19700978  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Yue, H., Thornton, M., Runkumar, J., Tang, Y.T., Azimzai, Y.,  
Baughn, M.R., Yang, J., Yao, M.G., Lal, P., Wallia, N.K., Gandhi, A.R.,  
Hafalia, A.J., Nguyen, D.B., Patterson, C., Elliott, V.S.,  
Tribouley, C.M., Lu, D.A., Xu, Y., Reddy, R., Hernandez, R.,  
Borowsky, M.L., Lo, T.P., Lu, Y., Policky, J.L., Greene, B.D.,  
Sanjanwala, M.S., Raumann, B.E., Burford, N., Ison, C.H., Lee, E.A.,  
Ding, L., Das, D., Kallick, D.A., Khan, P.A. and Seilhamer, J.J.  
Patent: WO 0212340-A 33 14-FEB-2002;  
Location/Qualifiers  
1..6369  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 2798241CBI"

Query Match 93.8%; Score 5135.6; DB 6; Length 6369;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5138; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 97 AGGTTTATTCAGAAACATGTCCTCAATGAGGAGTAGGAGTTGGAGACAGACCA 156  
DB 1221 AGGTTTATTCAGAAACATGTCCTCAATGAGGAGTAGGAGTTGGAGACAGACCA 1280  
QY 157 GAACACTTCTACTGAAGAATTACTTAATAATGAGAACCAAAAGAGTAGTGTTCAGG 216  
DB 1281 GAACACTTCTACTGAAGAATTACTTAATAATGAGAACCAAAAGAGTAGTGTTCAGG 1340  
QY 217 AAATCTTTTCCACTATTTTATTTTGGTTAAATTAATTAATGAGATGATGATCCAA 276  
DB 1341 AAATCTTTTCCACTATTTTATTTTGGTTAAATTAATTAATGAGATGATGATCCAA 1400  
QY 277 ATAAGAAATATGAGAAAGTGCTAATATAGAACTCAATCTATGGAAGTTACTCTTT 336  
DB 1401 ATAAGAAATATGAGAAAGTGCTAATATAGAACTCAATCTATGGAAGTTACTCTTT 1460  
QY 337 CTAATCTAATCTTGGATATCTCCAGTGACTAATATTAACAGCAGCATCATGCGAAG 396  
DB 1461 CTAATCTAATCTTGGATATCTCCAGTGACTAATATTAACAGCAGCATCATGCGAAG 1520  
QY 397 TGCTACTGATCATCTACCTGATGTCATPAATTAAGAAATATACAAATGAAAGAA 456  
DB 1521 TGCTACTGATCATCTACCTGATGTCATPAATTAAGAAATATACAAATGAAAGAA 1580  
QY 457 TGTTAACTCCAGTCTCTTAAGCGGACACTTTGAGTGTGGTTTTCAGACTCCA 516  
DB 1581 TGTTAACTCCAGTCTCTTAAGCGGACACTTTGAGTGTGGTTTTCAGACTCCA 1640  
QY 517 TGCTCTATGAACCTCGTTTCTGATATGATTCAGTATCTTTCTATTTATATGGATT 576  
DB 1641 TGCTCTATGAACCTCGTTTCTGATATGATTCAGTATCTTTCTATTTATATGGATT 1700  
QY 577 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTGCTCAGGTTTCAG 636  
DB 1701 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTGCTCAGGTTTCAG 1760  
QY 637 TTTTCAAGCATCCATAGATGCTGCCATTATACAGTTTGAAGCAACAAATGTTCTTTGGA 696  
DB 1761 TTTTCAAGCATCCATAGATGCTGCCATTATACAGTTTGAAGCAACAAATGTTCTTTGGA 1820  
QY 697 AGGAGCTGGAGTCAACTAAAGCTGTTATTAATGAGAGAACTGCTGTTGTAGAAATAGATA 756

DB 1821 AGGAGCTGGAGTCAACTAAAGCTGTTTATTAATGAGGAGAAACTGCTGTTGTAGAAATAGATA 1880  
QY 757 CTTTTCCCGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTACCTTTTGGTACT 816  
DB 1881 CTTTTCCCGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTACCTTTTGGTACT 1940  
QY 817 TTTTGGCAATTCATATCGTAGCGAAGAAAGAAAAATTAAGAAATTTTAAAGATAA 876  
DB 1941 TTTTGGCAATTCATATCGTAGCGAAGAAAGAAAAATTAAGAAATTTTAAAGATAA 2000  
QY 877 TGGGCACTTCATGACTACTGCTTTTGGCTTTCTGGGTTCTTCTATATACAAAGTTTAATT 936  
DB 2001 TGGGCACTTCATGACTACTGCTTTTGGCTTTCTGGGTTCTTCTATATACAAAGTTTAATT 2060  
QY 937 TTCTTATGCTCCCTTCTTATGCGAGTCATTTGCGACAGCTTCTTTGTTTATTTCTCAAAGTA 996  
DB 2061 TTCTTATGCTCCCTTCTTATGCGAGTCATTTGCGACAGCTTCTTTGTTTATTTCTCAAAGTA 2120  
QY 997 GCAGCACTTGATATTTCTGCTTTTCTTTTCTTTATGGAATATCATCTGATTTTGTGCTT 1056  
DB 2121 GCAGCACTTGATATTTCTGCTTTTCTTTTCTTTATGGAATATCATCTGATTTTGTGCTT 2180  
QY 1057 TAATGCTGACACTCTTTTAAATAAATAAATCAAAATGATGGGAATAGTTGAATTTTGTGA 1116  
DB 2181 TAATGCTGACACTCTTTTAAATAAATAAATCAAAATGATGGGAATAGTTGAATTTTGTGA 2240  
QY 1117 CTGTGCTTTTGGATTTATTTGGCTTTATGATATCTCATAGAAAGTTTCCCAATCTT 1176  
DB 2241 CTGTGCTTTTGGATTTATTTGGCTTTATGATATCTCATAGAAAGTTTCCCAATCTT 2300  
QY 1177 TAGTGTGGCTTTTCTAGTCTCTTTCTGCTACTGTACTTTGTGATTTGATGTCACAGGTCA 1236  
DB 2301 TAGTGTGGCTTTTCTAGTCTCTTTCTGCTACTGTACTTTGTGATTTGATGTCACAGGTCA 2360  
QY 1237 TGCATTTAGAAAGTTTAAATGAAGTGCTTCATTTCAAAATTTGATGCGAGGCGCATATC 1296  
DB 2361 TGCATTTAGAAAGTTTAAATGAAGTGCTTCATTTCAAAATTTGATGCGAGGCGCATATC 2420  
QY 1297 CTCATAATTAACAATTAATCATGCTCACACTTAATAGTATATTTCTATGCTCTTTGGCTG 1356  
DB 2421 CTCATAATTAACAATTAATCATGCTCACACTTAATAGTATATTTCTATGCTCTTTGGCTG 2480  
QY 1357 TCTATCTTGATCAAGTCAATCCAGGGAAATTTGGCTTAGCGAGATCACTTTATATTTC 1416  
DB 2481 TCTATCTTGATCAAGTCAATCCAGGGAAATTTGGCTTAGCGAGATCACTTTATATTTC 2540  
QY 1417 TGAAGCTTTCATATTTGGTCAAAAGAGCAAAAGAAATTTAGAGAGTTATCAGAGGCGCAATG 1476  
DB 2541 TGAAGCTTTCATATTTGGTCAAAAGAGCAAAAGAAATTTAGAGAGTTATCAGAGGCGCAATG 2600  
QY 1477 TTAATGGAATATTTAGTTTATGAAATTAATGAGCCAGTTTCTTCAGAAATTTGTAGGAA 1536  
DB 2601 TTAATGGAATATTTAGTTTATGAAATTAATGAGCCAGTTTCTTCAGAAATTTGTAGGAA 2660  
QY 1537 AAGAGCCATAGAAATTTAGTGTGTTTCAAGACATACAGAAAGAGGTGAAATGTGG 1596  
DB 2661 AAGAGCCATAGAAATTTAGTGTGTTTCAAGACATACAGAAAGAGGTGAAATGTGG 2720  
QY 1597 AGGCTTTGAGAAATTTGTCAATTTGATATATATGAGGTCAGATTACTGCTTACTTTGGCC 1656  
DB 2721 AGGCTTTGAGAAATTTGTCAATTTGATATATATGAGGTCAGATTACTGCTTACTTTGGCC 2780  
QY 1657 ACAGTGGAAACAGAAAGTAGTACATTTGATGAATATTTTGTGAGTCTGCCACCTTCG 1716  
DB 2781 ACAGTGGAAACAGAAAGTAGTACATTTGATGAATATTTTGTGAGTCTGCCACCTTCG 2840  
QY 1717 ATGGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 1776  
DB 2841 ATGGGTTTGCATCTATATATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAA 2900  
QY 1777 GAAAAATGATTTGGCAATTTGTCCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAG 1836  
DB 2901 GAAAAATGATTTGGCAATTTGTCCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAG 2960

1837 AAAATTATCAATTTTGGCTTCAATCAAAAGGAGATACCGCCAAACAATATAATACAAAGAG 1896  
1896  
2961 AAAATTATCAATTTTGGCTTCAATCAAAAGGAGATACCGCCAAACAATATAATACAAAGAG 3020  
1897 TGCAGAGGTTTACTAGATTATAGACATGACAGACTATCAAGATTAACCAAGCTTAAATAAT 1956  
3021 TGCAGAGGTTTACTAGATTATAGACATGACAGACTATCAAGATTAACCAAGCTTAAATAAT 3080  
1957 TAAGTGGTGGTCAAAAAGAGCTGTCAATTAGGAATGCTGTCTTGGGAACCCAAAGA 2016  
3081 TAAGTGGTGGTCAAAAAGAGCTGTCAATTAGGAATGCTGTCTTGGGAACCCAAAGA 3140  
2017 TACTGTGCTAGATGAACCAAGCTGGAATGGAACCCCTGTCTCGACATATGTATGGA 2076  
3141 TACTGTGCTAGATGAACCAAGCTGGAATGGAACCCCTGTCTCGACATATGTATGGA 3200  
2077 ATCTTTTAAATAACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATG 2136  
3201 ATCTTTTAAATAACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATG 3260  
2137 AGCTGACATCTCTGCAGATAGGAAGCTGTGATATCAAGGAATGCTGAATGTGTG 2196  
3261 AGCTGACATCTCTGCAGATAGGAAGCTGTGATATCAAGGAATGCTGAATGTGTG 3320  
2197 GTTCTTCAATGTTCTTCAAAAAGTAAATGGGGGATCGGCTACCGCCTGAGCATGTACATAG 2256  
3321 GTTCTTCAATGTTCTTCAAAAAGTAAATGGGGGATCGGCTACCGCCTGAGCATGTACATAG 3380  
2257 ACAAAATATGTGCCACAGAAATCTCTTTCTTCACTGGTTAAACAAATATACCTGGAGCTA 2316  
3381 ACAAAATATGTGCCACAGAAATCTCTTTCTTCACTGGTTAAACAAATATACCTGGAGCTA 3440  
2317 CTTTATTAACAAGAAATGACCAACAACTTGTGTATAGCTTGGCTTTCAAGGACATGGACA 2376  
3441 CTTTATTAACAAGAAATGACCAACAACTTGTGTATAGCTTGGCTTTCAAGGACATGGACA 3500  
2377 AATTTTCAGTTTGTGTTCTGCGCTAGACAGTCAATCAAAATTTGGGTGTCATCTTATG 2436  
3501 AATTTTCAGTTTGTGTTCTGCGCTAGACAGTCAATCAAAATTTGGGTGTCATCTTATG 3560  
2437 GTCTTTCCATGACCACTTTGGGAAGAGTATTTTAAAGCTAGAAGTTGAAGCAGAAATG 2496  
3561 GTCTTTCCATGACCACTTTGGGAAGAGTATTTTAAAGCTAGAAGTTGAAGCAGAAATG 3620  
2497 ACCAAGCAGATTATAGTATTTACTCAGAGCCACTGAGGAGAAATGGATCAAAAT 2556  
3621 ACCAAGCAGATTATAGTATTTACTCAGAGCCACTGAGGAGAAATGGATCAAAAT 3680  
2557 CTTTTCATGAATGGAAACAGAGCTTACTTATTTCTTGAACCAAGGCTTCTTAGTGA 2616  
3681 CTTTTCATGAATGGAAACAGAGCTTACTTATTTCTTGAACCAAGGCTTCTTAGTGA 3740  
2617 GCACCATGAGCCTTTGGAAACCAACAGATGTATACAAATAGCAAAATTTCAATTTCTTACCT 2676  
3741 GCACCATGAGCCTTTGGAAACCAACAGATGTATACAAATAGCAAAATTTCAATTTCTTACCT 3800  
2677 TGAACGCTGAAGTAATACAGTGAGATCAGTGTGCTTCTGCTTCTTAAATTTTTCACAG 2736  
3801 TGAACGCTGAAGTAATACAGTGAGATCAGTGTGCTTCTGCTTCTTAAATTTTTCACAG 3860  
2737 TGCAGATTTTATGTTTTTGGTTTCATCACTCTTTTAAATAAGCTGTGGTTCCCATCAAC 2796  
3861 TGCAGATTTTATGTTTTTGGTTTCATCACTCTTTTAAATAAGCTGTGGTTCCCATCAAC 3920  
2797 TTTGTTCCAGACTTATTTTCTTAAACCTGGAGACAAACCAATATACAAACAGTCT 2856  
3921 TTTGTTCCAGACTTATTTTCTTAAACCTGGAGACAAACCAATATATACAAACAGTCT 3980  
2857 TGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 2916  
3981 TGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 4040

QY 2917 AGAACATAATGTTGTCAGCATGATTAATGACAGTGACTATGATCCGTGGCTCCCATAGTG 2976  
Db 4041 AGAACATAATGTTGTCAGCATGATTAATGACAGTGACTATGATCCGTGGCTCCCATAGTG 4100  
QY 2977 CGGCTTTAAATGTGATGCACTTCAAGAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTA 3036  
Db 4101 CGGCTTTAAATGTGATGCACTTCAAGAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTA 4160  
QY 3037 CTATGTTTAAATCTTTACCTATATTAAGTGAATATCATTAAGTAACTACTATCTTTATCAAT 3096  
Db 4161 CTATGTTTAAATCTTTACCTATATTAAGTGAATATCATTAAGTAACTACTATCTTTATCAAT 4220  
QY 3097 TAAATGTGACTGAAAAACCATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTAAGTGA 3156  
Db 4221 TAAATGTGACTGAAAAACCATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTAAGTGA 4280  
QY 3157 TAGTTTTTAAATTTAGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTTACGAA 3216  
Db 4281 TAGTTTTTAAATTTAGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTTACGAA 4340  
QY 3217 TGCACACTTACTTTGCCATGGAATAATGCGAGAGATCATTAAGTCAAAAGCTTATACCTAAC 3276  
Db 4341 TGCACACTTACTTTGCCATGGAATAATGCGAGAGATCATTAAGTCAAAAGCTTATACCTAAC 4400  
QY 3277 TTAACCTTTCAAGTCTTTTGGCCATCTGCAATTTGGATTTGGACAAGCTGTTTGTATATCC 3336  
Db 4401 TTAACCTTTCAAGTCTTTTGGCCATCTGCAATTTGGATTTGGACAAGCTGTTTGTATATCC 4460  
QY 3337 CCTTATTTTATCATCTTATTTTGCATGCTAGGAAGCTTATGGCAATTCATTTATGGAT 3396  
Db 4461 CCTTATTTTATCATCTTATTTTGCATGCTAGGAAGCTTATGGCAATTCATTTATGGAT 4520  
QY 3397 TATATTTTATCTGTAAGAGTCTCTTGTGCTGTTTTTGGCTTATTTGGTATGTTCCAT 3456  
Db 4521 TATATTTTATCTGTAAGAGTCTCTTGTGCTGTTTTTGGCTTATTTGGTATGTTCCAT 4580  
QY 3457 CAGTATTTCTGTTCACTTATTTTCTTCTTCACTTTAAGAAATTTTAAATACCAAG 3516  
Db 4581 CAGTATTTCTGTTCACTTATTTTCTTCTTCACTTTAAGAAATTTTAAATACCAAG 4640  
QY 3517 AATTTTGGTCAATTTATCTATTTCTGTCGAGCGTTGCTTGTATTTGCAATCACTGAAATAA 3576  
Db 4641 AATTTTGGTCAATTTATCTATTTCTGTCGAGCGTTGCTTGTATTTGCAATCACTGAAATAA 4700  
QY 3577 CTTTCTTTATGGGATACAGAAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCTTC 3636  
Db 4701 CTTTCTTTATGGGATACAGAAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCTTC 4760  
QY 3637 CAATCTATCCACTTTAGGTTGCTGATTTCTTCAATAAGATTTCTTGGAGAATGTAC 3696  
Db 4761 CAATCTATCCACTTTAGGTTGCTGATTTCTTCAATAAGATTTCTTGGAGAATGTAC 4820  
QY 3697 GAAAAATGGAACACTATAATCCATGGGATAGGCTTTTCACTAGCTGTTTATATCGCCTT 3756  
Db 4821 GAAAAATGGAACACTATAATCCATGGGATAGGCTTTTCACTAGCTGTTTATATCGCCTT 4880  
QY 3757 ACCTGCACTGTGTACTGTGATTTTCTCTTCACTATGAGAAAAATATGGAGGCA 3816  
Db 4881 ACCTGCACTGTGTACTGTGATTTTCTCTTCACTATGAGAAAAATATGGAGGCA 4940  
QY 3817 GATCAATTAAGAAAGATTCCTTTTTCAGAAACCTTTTCAACGAAAGTCTAAAAATAGGAAGC 3876  
Db 4941 GATCAATTAAGAAAGATTCCTTTTTCAGAAACCTTTTCAACGAAAGTCTAAAAATAGGAAGC 5000  
QY 3877 TTTCCAGAACCCAGACCAATGAGATGAAGATGTCAAGCTGAAAGCTAAAGG 3936  
Db 5001 TTTCCAGAACCCAGACCAATGAGATGAAGATGTCAAGCTGAAAGCTAAAGG 5060  
QY 3937 TCAAGAGCTGATGGGTGCCAGTGTGTGAGGAGAAACCATTCATATAGTGTGAGCAAT 3996  
Db 5061 TCAAGAGCTGATGGGTGCCAGTGTGTGAGGAGAAACCATTCATATAGTGTGAGCAAT 5120  
QY 3997 TGCATAAAGATATATGACAAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGG 4056

```
Db 5121 TGCATTAAGAAATATGATGACCAAGAAAGATTTCTTTCTTCAAGAAAGTAAAGAAAGTGG 5180
Qy 4057 CAACTAAATACATCTCTTTCTGTGTGAAAGAGAGAGATCTTAGGACTATTGGGTCCAA 4116
Db 5181 CAACTAAATACATCTCTTTCTGTGTGAAAGAGAGATCTTAGGACTATTGGGTCCAA 5240
Qy 4117 ATGGTGTGGCAAGACCAATATATATATTTCTGGTTGGTGATATTGAACCACTTCAG 4176
Db 5241 ATGGTGTGGCAAGACCAATATATATATTTCTGGTTGGTGATATTGAACCACTTCAG 5300
Qy 4177 GCCAGTATTTTAGGAGATTTCTTTCAGAGCAAGTGAAGATGATGATTCACCTGAAGT 4236
Db 5301 GCCAGTATTTTAGGAGATTTCTTTCAGAGCAAGTGAAGATGATGATTCACCTGAAGT 5360
Qy 4237 GTATGGTTACTGTCTCAGATAAACCCCTTGTGGCCAGATACATACATTGAGGAACATT 4296
Db 5361 GTATGGTTACTGTCTCAGATAAACCCCTTGTGGCCAGATACATACATTGAGGAACATT 5420
Qy 4297 TTGAATTTATGAGCTGTCAAGGAATGAGTGAAGTGAATGAAGAACTCATAGTC 4356
Db 5421 TTGAATTTATGAGCTGTCAAGGAATGAGTGAAGTGAATGAAGAACTCATAGTC 5480
Qy 4357 GAATAACACATCCTTGATTTAAAGAAACATCTTCAGAAGACTGTAAAGAACTACCTG 4416
Db 5481 GAATAACACATCCTTGATTTAAAGAAACATCTTCAGAAGACTGTAAAGAACTACCTG 5540
Qy 4417 CAGGAATCAAGAAAGTTGTCTTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTT 4476
Db 5541 CAGGAATCAAGAAAGTTGTCTTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTT 5600
Qy 4477 TGTAGATGAACCATCTCAGATGATGATCCAAAGCCAAACAGACATGTGGCGAGCAA 4536
Db 5601 TGTAGATGAACCATCTCAGATGATGATCCAAAGCCAAACAGACATGTGGCGAGCAA 5660
Qy 4537 TTGCACTGCATTTAAACAGAAAGCGGCTGCTATTCTGACCACTCCTATATGAGG 4596
Db 5661 TTGCACTGCATTTAAACAGAAAGCGGCTGCTATTCTGACCACTCCTATATGAGG 5720
Qy 4597 AGGCAGAGCTGTCTGTGATCAGTAGTATCATGTGCTCTGGGAGTTAAGATGATCG 4656
Db 5721 AGGCAGAGCTGTCTGTGATCAGTAGTATCATGTGCTCTGGGAGTTAAGATGATCG 5780
Qy 4657 GAACAGTACAAATCAAGAGTAAATTTGAAAGAGCTACTTTTGGAAATTTAAATGA 4716
Db 5781 GAACAGTACAAATCAAGAGTAAATTTGAAAGAGCTACTTTTGGAAATTTAAATGA 5840
Qy 4717 AGGACTGGATAGAAACCTTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTCC 4776
Db 5841 AGGACTGGATAGAAACCTTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTCC 5900
Qy 4777 CAAATGCAAGCGCTCAGGAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAG 4836
Db 5901 CAAATGCAAGCGCTCAGGAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAG 5960
Qy 4837 ATGTTCAAGTCCCTTTTCAAACTTTTATAGCTGGAAGAGCTTAAACATGCTTTGCCA 4896
Db 5961 ATGTTCAAGTCCCTTTTCAAACTTTTATAGCTGGAAGAGCTTAAACATGCTTTGCCA 6020
Qy 4897 TTGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGTTTGTAGAACTCACTAAG 4956
Db 6021 TTGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGTTTGTAGAACTCACTAAG 6080
Qy 4957 AACAGAGGAGGAAGATATAGTTGTGGAACCTTTTAAACAGACACATTTTGTGGGAAGCAA 5016
Db 6081 AACAGAGGAGGAAGATATAGTTGTGGAACCTTTTAAACAGACACATTTTGTGGGAAGCAA 6140
Qy 5017 CACAAGAGATAGCTAGTATTTTGAATTTGTATTTTGTGCTGCTTACTTGGGACTTCT 5076
Db 6141 CACAAGAGATAGCTAGTATTTTGAATTTGTATTTTGTGCTGCTTACTTGGGACTTCT 6200
Qy 5077 TTCTTTTCACTAATTTTAACTTTTGTGTTTAAAGTTTTTATTTGAATGCTTAACCTGA 5136
```

```
Db 6201 TTCTTTTCACTAATTTTAACTTTGTTTAAAAAGTTTTTTTATGGAATGGAATCTGA 6260
Qy 5137 GAACCAAGAACGACACTTGAATTTTCTTAAGCTCCCTTAATTTGAAATGCTGCTGTTGTG 5196
Db 6261 GAACCAAGAACGACACTTGAATTTTCTTAAGCTCCCTTAATTTGAAATGCTGCTGTTGTG 6320
Qy 5197 TTTTGTCTTTCTTTAAATAAAACGATGATGATATTAAGTGA 5238
Db 6321 TTTTGTCTTTCTTTAAATAAAACGATGATGATATTAAGTGA 6362

RESULT 6
HSA275973 5096 bp mRNA linear PRI 31-JUL-2002
LOCUS Homo sapiens mRNA for ATP-binding cassette protein of the (ABCA
DEFINITION subfamily).
ACCESSION AJ275973
VERSION AJ275973.3 GI:22080663
KEYWORDS ABCA subfamily; ATP-binding cassette protein.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Petry, F., Kotthaus, A. and Hirsch-Ernst, K.I.
TITLE Molecular cloning and tissue distribution of a novel ATP-binding
cassette (ABC) transporter belonging to the subfamily ABCA
JOURNAL Unpublished
REFERENCE 2
AUTHORS Hirsch-Ernst, K.I.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2000) Hirsch-Ernst K.I., Department of
Toxicology, University of Goettingen, Robert-Koch-Strasse 40,
Lowery Saxony, D-37075, GERMANY
REMARK revised by [3]
REFERENCE 3 (bases 1 to 5096)
AUTHORS Hirsch-Ernst, K.I.
TITLE Direct Submission
JOURNAL Submitted (20-NOV-2001) Hirsch-Ernst K.I., Department of
Toxicology, University of Goettingen, Robert-Koch-Strasse 40,
Lowery Saxony, D-37075, GERMANY
COMMENT On Aug 1, 2002 this sequence version replaced gi:17046099.
FEATURES
source
1..5096
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="testis"
66..4994
/function="ABC transporter"
/notes="belongs to the ABCA subfamily"
/codon_start=1
/product="ATP-binding cassette protein"
/protein_id="CA93535.3"
/db_xref="GI:22080664"
/db_xref="GOA:Q9NY14"
/db_xref="SPTREMBL:Q9NY14"
/translation="MSTAIREVGVWRQRTLLKLYLKRCRKSSVQELPPLPLFLP
WILISMMHPNKKYBEVNPINELNPMDFKTLNLIILGYTPVTNITSIIMQKVDHLDP
VITTEYTNKEKMLTSSLSKPSNFVGVFKDSMSYELRFPDPMI PVSITMDSRAGCS
KSCAAQWYSGFTVLOASIDAIIQLKTNVSLWKELESTKAVIMGETAVVIDTPPR
GVLLIVLAFSPFGVFLAIHVAEKKIKKEFLKIMGLHDTAFWL SWLLYLSLIFL
MSLLMVAINTASLLPQSSSIVIFLLFELYGLSSVFPFALMTPLPKSKHVGIVRFFV
TVAFGIUMILIESFPASLVWLPSPFCFTFVIGAQVMELEDFNEGASPSNLTAG
PYPLITITMLTNSIFVLLAVLDQVIFGPFGLRRSLYFLKPSYKSRNRYEEL
SEGNVNGNISFSEIIEPVSEFVGKEAIRISGIQTKYRKKGNEVALRNLSPDIYEGQ
ITALLCHSGTSGTLMNLCIGCPPSDGPFASITGHRVSEIDEMFARKMIGICPDII
HFDVLTVENLSIASIKGI PANNIIOEVQKVLDDMOTIKDNOAKKLSGGQKRKLS
LGIVLGNPKIILLDEPTAGMDPCSRHIVNLLKYRKANRVTVFSHFVDEADILADR
KAVISGGMKLCVGSNNFLKSKWIGIGYRLSWYDKYCATESLSL VKOHTPGATLQON
DQQLVYSDPDKMDKPSGFSALDSHNLGVISYGSMTLLEDFVFLKLEVEAEIDQAD
YSVFTQPLEEEMDSKDFEMLISLETSKASLSVSTNLMKQKMTIAKHFHFFLK
RESKSVRSVRLILLLIFFTVQIFMFLVHHSFKNAVVPILKLPDPLKPKGDKPHKYKTS
```



LLQNSADSDISDLISPTSONIMVWINDSDYVAVPHSAALNVHSEKDYVFAVAF  
NSTWVYSLPILNIIISNYLYHLNVETIOIWSPTFQEIITDIFKIELYFOALIGI  
IVTAMPYFAMENAEHKKAYTOQLKSLGSPAYMIGQAVDIPFIIILMLGSL  
LAFHGYLFYVVKFLAVVCLIGVPSVILFTVIASTFKKILNLTKEFMSFIYSVAAL  
ACIAITBITFMGTIATILHYAPCIIIPVILLGCLISPIKISWKNVRKNVNTYPMW  
DRLSVANTSPVLOCLMIFLOVYKVKYKGRSIRKDPFFENLSTKSKNPKLPPBPNE  
DEDEDVKAERLYELWGCQCCCEPSIMVSNLHKEYDDKDFILLRKYKVKYATKIS  
FCVKGELGLLGNAGKSTIINILVGDIEPTSGQIFLGDYSESEDDSDSLKMGY  
CPQINLPWFDITLQHEFEIYAGVKGMSDMKEVISIRITHALDLKHLQKTVKGLGAG  
IKRKLCTALSMGLNGPQLIDBEPSTGMDPKAKHMRRAIRTAFAKRNKRAAILTYHME  
EABAVCORVAIMVSGQICITGVQHLKSKFKGKVFLEIKLMDWIEINLEVDRLQREIQY  
IFPNASRQESPSLAIKPKEDVQSLQSFKKLEAKHAFATIEEYSFSQATLEQVVF  
ELTKEOEEDNSCOTLNSLWERTQEDRVF"

## ORIGIN

Query Match 92.8%; Score 5083.2; DB 9; Length 5096;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 5085; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	57	GAGCGCTCCAGGCTGACAGCTCTGGCGCTCGGGCCCTGAGGTTTATTTCAGAAAACATG	116
DB	9	GAGCGCTCCAGGCTGACAGCTCTGGCGCTCGGGCCCTGAGGTTTATTTCAGAAAACATG	68
QY	117	TCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCAAGCACTTCTACTCAAGAAAT	176
DB	69	TCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCAAGCACTTCTACTCAAGAAAT	128
QY	177	TACTTAATTAATCCAGAACCAAAAGAGTAGTCTTCAGGAAATCTTTTTTCCACTATTT	236
DB	129	TACTTAATTAATCCAGAACCAAAAGAGTAGTCTTCAGGAAATCTTTTTTCCACTATTT	188
QY	237	TTTTTATTTTGGTTAAATTAATTAGCATGATGCATCCAAATAGAAAATATGAAAGAGTG	296
DB	189	TTTTTATTTTGGTTAAATTAATTAGCATGATGCATCCAAATAGAAAATATGAAAGAGTG	248
QY	297	CCTAATATAGAACTCAATCCTATGGAAGAGTTTACTCTTCTAATCTAATCTTGGATAT	356
DB	249	CCTAATATAGAACTCAATCCTATGGAAGAGTTTACTCTTCTAATCTAATCTTGGATAT	308
QY	357	ACTCCAGTGACTAATATTACAGCAGAGCATCATGACAGAAAGTGCTACTGATCATCTACCT	416
DB	309	ACTCCAGTGACTAATATTACAGCAGAGCATCATGACAGAAAGTGCTACTGATCATCTACCT	368
QY	417	GATGTCATAATTACTGAAGAATATACAAATGAAAAGAAATGTTAAACATCCAGTCTCTCT	476
DB	369	GATGTCATAATTACTGAAGAATATACAAATGAAAAGAAATGTTAAACATCCAGTCTCTCT	428
QY	477	AAGCGAGCAACTTTGTAGGTGTGGTTTCAAGACCTCCATGCTCTATGAACCTTGGTTT	536
DB	429	AAGCGAGCAACTTTGTAGGTGTGGTTTCAAGACCTCCATGCTCTATGAACCTTGGTTT	488
QY	537	TTTCTGATATGATTCAGATATCTTCTATTTATATAGGATTCAGAGCTGGCTGTTCAAA	596
DB	489	TTTCTGATATGATTCAGATATCTTCTATTTATATAGGATTCAGAGCTGGCTGTTCAAA	548
QY	597	TGATGTAGGCTGCTCAGTACTGGTCTCAGTTTTCAGTTTTCAGAGTCCATAGAT	656
DB	549	TGATGTAGGCTGCTCAGTACTGGTCTCAGTTTTCAGTTTTCAGAGTCCATAGAT	608
QY	657	GCTGCCATATACAGTTGAAGACCAATGTTTCTCTTTGGAGAGAGCTGGAGTCACTAAA	716
DB	609	GCTGCCATATACAGTTGAAGACCAATGTTTCTCTTTGGAGAGAGCTGGAGTCACTAAA	668
QY	717	GCTGTTATTATGGGAGAACTGCTGTTGTGTAAGAAATAGATACCTTTCCCGAGGAGTAAT	776
DB	669	GCTGTTATTATGGGAGAACTGCTGTTGTGTAAGAAATAGATACCTTTCCCGAGGAGTAAT	728
QY	777	TTAATATACCTAGTTATAGCAATTTTCACTTTTGGATATTTTGGCAATTCATATCGTA	836
DB	729	TTAATATACCTAGTTATAGCAATTTTCACTTTTGGATATTTTGGCAATTCATATCGTA	788
QY	837	CGAGAAAGAGAAAATAAAGAAATTTTAAAGATAATGGACATTCATGATCTGCC	896

DB	789	GCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACTGCC	848
QY	897	TTTTGGCTTTCTCGGGTTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCTTCTTATG	956
DB	849	TTTTGGCTTTCTCGGGTTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCTTCTTATG	908
QY	957	GCAGTCATTGGACAGCTCTTTTGTATTTCCTCAAGTAGCAGCAATTTGTGATATTCTG	1016
DB	909	GCAGTCATTGGACAGCTCTTTTGTATTTCCTCAAGTAGCAGCAATTTGTGATATTCTG	968
QY	1017	CTTTTCTTCTTATAGGATATCATCTGATTTTGTCTTAAATCTGACACCTCTTTT	1076
DB	969	CTTTTCTTCTTATAGGATATCATCTGATTTTGTCTTAAATCTGACACCTCTTTT	1028
QY	1077	AAAAAATCAAAACATGTGGGAATAGTTGAAATTTTGTACTGTGGCTTTTGGATTTAT	1136
DB	1029	AAAAAATCAAAACATGTGGGAATAGTTGAAATTTTGTACTGTGGCTTTTGGATTTAT	1088
QY	1137	GGCTTTATGATAATCCTCATAGAAAGTTTCCAAATCGTTAGTGTGGCTTTTCACTCCT	1196
DB	1089	GGCTTTATGATAATCCTCATAGAAAGTTTCCAAATCGTTAGTGTGGCTTTTCACTCCT	1148
QY	1197	TTCTGTCACTGTACTTTTGTGATTTGATTTGACAGGTTCATGCAATTTAGAGATTTTAA	1256
DB	1149	TTCTGTCACTGTACTTTTGTGATTTGATTTGACAGGTTCATGCAATTTAGAGATTTTAA	1208
QY	1257	GAAGTGTCTTCAATTTTCAAAATTTGACTGCAGGCCCATATCCTTAATTTTACAATATTC	1316
DB	1209	GAAGTGTCTTCAATTTTCAAAATTTGACTGCAGGCCCATATCCTTAATTTTACAATATTC	1268
QY	1317	ATGCTCACACTTAATAGTATATTTATGCTCTTTGGCTGTCTATCTTGTGATCAAGTCAT	1376
DB	1269	ATGCTCACACTTAATAGTATATTTATGCTCTTTGGCTGTCTATCTTGTGATCAAGTCAT	1328
QY	1377	CCAGGGAAATTTGGCTTACGGAGATCATCTTATATTTTCTGAAGCTTCATATTTGGTCA	1436
DB	1329	CCAGGGAAATTTGGCTTACGGAGATCATCTTATATTTTCTGAAGCTTCATATTTGGTCA	1388
QY	1437	AAGAGCAAAAGAAATTTAGAGAGTTTATCAGAGGCAATTTAAATGGAAATTTAGTTTT	1496
DB	1389	AAGAGCAAAAGAAATTTAGAGAGTTTATCAGAGGCAATTTAAATGGAAATTTAGTTTT	1448
QY	1497	AGTGAATTTTGGCCAGTTTCTTCAGAAATTTGAGAAAGAGCCATAGAAATTTAGT	1556
DB	1449	AGTGAATTTTGGCCAGTTTCTTCAGAAATTTGAGAAAGAGCCATAGAAATTTAGT	1508
QY	1557	GATTTTCAGAGACATACAGAAAGAGGTGAAATTTGGAGGCTTTGAGAAATTTGTC	1616
DB	1509	GATTTTCAGAGACATACAGAAAGAGGTGAAATTTGGAGGCTTTGAGAAATTTGTC	1568
QY	1617	TTTGACATATGAGGTCAGATTTACTGCTTACTTGGCCACAGTGGAACAGGAAGAGT	1676
DB	1569	TTTGACATATGAGGTCAGATTTACTGCTTACTTGGCCACAGTGGAACAGGAAGAGT	1628
QY	1677	ACATTTGATGAATTTCTTTTGTGACTCTGCCACCTCTCTGATGGTTTGCATCTATATAT	1736
DB	1629	ACATTTGATGAATTTCTTTTGTGACTCTGCCACCTCTCTGATGGTTTGCATCTATATAT	1688
QY	1737	GGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAAGAAATGATTTGGCAATTTGT	1796
DB	1689	GGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAAGAAATGATTTGGCAATTTGT	1748
QY	1797	CCACAGTTAGATATACATCTTTGATTTTTCAGTAGAAGAAATTTTATCAATTTTGGCT	1856
DB	1749	CCACAGTTAGATATACATCTTTGATTTTTCAGTAGAAGAAATTTTATCAATTTTGGCT	1808
QY	1857	TCAATCAAGGATACACAGCCCAATATATATACAGAGGTGCAGAGGTTTACTAGAT	1916
DB	1809	TCAATCAAGGATACACAGCCCAATATATATACAGAGGTGCAGAGGTTTACTAGAT	1868
QY	1917	TTAGACATCCAGCATATCAAGATAACCAAGCTTAAAAATTTAAGTGGTGGTCAAAAAAGA	1976
DB	1869	TTAGACATCCAGCATATCAAGATAACCAAGCTTAAAAATTTAAGTGGTGGTCAAAAAAGA	1928



QY 1977 AAGCTGTCATTAGGAATTCCTTGGGAACCCAAAGATACCTGCTAGATGACCA 2036  
Db 1929 AAGCTGTCATTAGGAATTCCTTGGGAACCCAAAGATACCTGCTAGATGACCA 1988  
QY 2037 ACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATACAGAAAA 2096  
Db 1989 ACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATACAGAAAA 2048  
QY 2097 GCCAATCGGTGACAGTGTCTCAGTACTCATTTTCATGATGAGTGAAGCTCATTTTCGAGAT 2156  
Db 2049 GCCAATCGGTGACAGTGTCTCAGTACTCATTTTCATGATGAGTGAAGCTCATTTTCGAGAT 2108  
QY 2157 AGGAAAGCTGTGATATCAAAAGGAATGCTGAAATGTGTGTTCTTCAATGTTCCTCAAA 2216  
Db 2109 AGGAAAGCTGTGATATCAAAAGGAATGCTGAAATGTGTGTTCTTCAATGTTCCTCAAA 2168  
QY 2217 AGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAATAATTTGCGCACAGAA 2276  
Db 2169 AGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAATAATTTGCGCACAGAA 2228  
QY 2277 TCTCTTTCTTCACTGGTTTAAACAACATATACCTCGAGCTACTTTTATTAACAAGATGAC 2336  
Db 2229 TCTCTTTCTTCACTGGTTTAAACAACATATACCTCGAGCTACTTTTATTAACAAGATGAC 2288  
QY 2337 CAACAACCTGTGTATAGCTTGCCTTTCAAGACATGACAAAATTTTCAGGTTTGTCTTCT 2396  
Db 2289 CAACAACCTGTGTATAGCTTGCCTTTCAAGACATGACAAAATTTTCAGGTTTGTCTTCT 2348  
QY 2397 GCCCTAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTATGTGTGTTCCTCATGACACTTG 2456  
Db 2349 GCCCTAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTATGTGTGTTCCTCATGACACTTG 2408  
QY 2457 GAAGACGTTATTTTAAAGCTAGAAAGTGAAGACAGAAATGACCAAGCAGATTAATAGTGA 2516  
Db 2409 GAAGACGTTATTTTAAAGCTAGAAAGTGAAGACAGAAATGACCAAGCAGATTAATAGTGA 2468  
QY 2517 TTTTACTCAGACGACCTCGAGGAGAAATGAAATCAAAATCTTTTGTATGAATGGACAG 2576  
Db 2469 TTTTACTCAGACGACCTCGAGGAGAAATGAAATCAAAATCTTTTGTATGAATGGACAG 2528  
QY 2577 AGCTTACTTATTTCTTCTGAACCAAGCTTCTCTAGTGAGCACCATGAGCCCTTTGGAA 2636  
Db 2529 AGCTTACTTATTTCTTCTGAACCAAGCTTCTCTAGTGAGCACCATGAGCCCTTTGGAA 2588  
QY 2637 CAACAGATGTATACAAATAGCAAAAGTTTCAATTTCTTTACCTTGAAACGTGAAGTAATCA 2696  
Db 2589 CAACAGATGTATACAAATAGCAAAAGTTTCAATTTCTTTACCTTGAAACGTGAAGTAATCA 2648  
QY 2697 GTGAGATCAGTGTGTCTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTG 2756  
Db 2649 GTGAGATCAGTGTGTCTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTG 2708  
QY 2757 GTTCATGACTCTTTTAAATGCTGTGGTTCCTCATCAAACTGTGTTCCAGACTTATATTTT 2816  
Db 2709 GTTCATGACTCTTTTAAATGCTGTGGTTCCTCATCAAACTGTGTTCCAGACTTATATTTT 2768  
QY 2817 CTAAAACTGAGACAAACCAATATAAACAAGTCTGCTTCTTTCAAAATTTCTGCT 2876  
Db 2769 CTAAAACTGAGACAAACCAATATAAACAAGTCTGCTTCTTTCAAAATTTCTGCT 2828  
QY 2877 GACTCAGATATCAGTGTATTTATGATGTTTTCACAGCCAGAACATAATGTTGAGCATG 2936  
Db 2829 GACTCAGATATCAGTGTATTTATGATGTTTTCACAGCCAGAACATAATGTTGAGCATG 2888  
QY 2937 ATTAATGACGACTATGATCGTGCCTCCCATAGTCGGCTTTTAAATGTTGATGAT 2996  
Db 2889 ATTAATGACGACTATGATCGTGCCTCCCATAGTCGGCTTTTAAATGTTGATGAT 2948  
QY 2997 TCAGAAAAGGACTATGTTTTTGCAGCTGTTTTCAACAGTACTATGGTTTTATTTTACCT 3056  
Db 2949 TCAGAAAAGGACTATGTTTTTGCAGCTGTTTTCAACAGTACTATGGTTTTATTTTACCT 3008

QY 3057 ATATTAGTGAATATCATTTAGTAACTACTATCTTTTATCATTTTAAATGTGACTGAAACCATC 3116  
Db 3009 ATATTAGTGAATATCATTTAGTAACTACTATCTTTTATCATTTTAAATGTGACTGAAACCATC 3068  
QY 3117 CAGATCTGAGTACCCCAATCTTTTCAAGAAATTAAGTATATAGTTTAAAAATTTGAGCTG 3176  
Db 3069 CAGATCTGAGTACCCCAATCTTTTCAAGAAATTAAGTATATAGTTTAAAAATTTGAGCTG 3128  
QY 3177 TATTTTCAAGCAGCTTTCCTTGGAAATCAATGTTTACTGCAATGCGCACTTACTTTGCCATG 3236  
Db 3129 TATTTTCAAGCAGCTTTCCTTGGAAATCAATGTTTACTGCAATGCGCACTTACTTTGCCATG 3188  
QY 3237 GAAATGACAGAGAATCAATAAGATCAAGCTTATATCTCAACTTTAACTTTTCAGGCTTTTG 3296  
Db 3189 GAAATGACAGAGAATCAATAAGATCAAGCTTATATCTCAACTTTAACTTTTCAGGCTTTTG 3248  
QY 3297 CCATCTGCATATTTGATGAGTGGCAAGCTGTTGTTGATATCCCTTATTTTATCATTTCT 3356  
Db 3249 CCATCTGCATATTTGATGAGTGGCAAGCTGTTGTTGATATCCCTTATTTTATCATTTCT 3308  
QY 3357 ATTTTGTAGCTAGGAAGCTTATTTGGCATTTTCAATATGGAATATATTTTATACTGTAAG 3416  
Db 3309 ATTTTGTAGCTAGGAAGCTTATTTGGCATTTTCAATATGGAATATATTTTATACTGTAAG 3368  
QY 3417 TTCCTTGTGCTGTTTTTTCCTTATTTGTTTATGTTTCCATCAGTTATTTCTGTTCACTTAT 3476  
Db 3369 TTCCTTGTGCTGTTTTTTCCTTATTTGTTTATGTTTCCATCAGTTATTTCTGTTCACTTAT 3428  
QY 3477 ATTGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTCTAT 3536  
Db 3429 ATTGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTCTAT 3488  
QY 3537 TCTGTGGCAGCTTGGCTTGTATTCGAATCACTGCAATTAACCTTTCTTTATGGGATACACA 3596  
Db 3489 TCTGTGGCAGCTTGGCTTGTATTCGAATCACTGCAATTAACCTTTCTTTATGGGATACACA 3548  
QY 3597 ATTGCAACTTATTTCTTATTTGCTTTTGTATTCATTTCCAAATCTTCCACTCTCACTGAGT 3656  
Db 3549 ATTGCAACTTATTTCTTATTTGCTTTTGTATTCATTTCCAAATCTTCCACTCTCACTGAGT 3608  
QY 3657 TGCCTGATTTCTTTCATAAAGATTTTCTTGGAAAGATGTACGAAATAATTTGTGACACCTAT 3716  
Db 3609 TGCCTGATTTCTTTCATAAAGATTTTCTTGGAAAGATGTACGAAATAATTTGTGACACCTAT 3668  
QY 3717 AATCCATGGGATAGCTTTTTCAGTAGCTTTATATCGCTTACCTGCTGAGTGTACTGTGG 3776  
Db 3669 AATCCATGGGATAGCTTTTTCAGTAGCTTTATATCGCTTACCTGCTGAGTGTACTGTGG 3728  
QY 3777 ATTTTCTCTTACAATACTATGAGAAAAATATGGAGGCAGATCAATTAAGAAAAAGATCCC 3836  
Db 3729 ATTTTCTCTTACAATACTATGAGAAAAATATGGAGGCAGATCAATTAAGAAAAAGATCCC 3788  
QY 3837 TTTTTCAGAAACCTTTTCAACGAGTCTTAAATAATAGGAAGCTTCCAGAACCCAGACAAAT 3896  
Db 3789 TTTTTCAGAAACCTTTTCAACGAGTCTTAAATAATAGGAAGCTTCCAGAACCCAGACAAAT 3848  
QY 3897 GAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTTAAAGGTCAAAGAGCTGATGGTTGC 3956  
Db 3849 GAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTTAAAGGTCAAAGAGCTGATGGTTGC 3908  
QY 3957 CAGTGTGTGAGGAAACCATTCATTTATGTTGAGCAATTTGGCATTAAGAAATATGATGAC 4016  
Db 3909 CAGTGTGTGAGGAAACCATTCATTTATGTTGAGCAATTTGGCATTAAGAAATATGATGAC 3968  
QY 4017 AAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTGGCAACTTAAATACATCTCTTTTC 4076  
Db 3969 AAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTGGCAACTTAAATACATCTCTTTTC 4028  
QY 4077 TGTGTGAAAAAGGAGAGATCTTTAGGACTATTTGGGTCCAAATGGTGTCTGGCAAAAGCACA 4136  
Db 4029 TGTGTGAAAAAGGAGAGATCTTTAGGACTATTTGGGTCCAAATGGTGTCTGGCAAAAGCACA 4088  
QY 4137 ATTTATTAATATCTGTTGGTGTATTTGAACCAACTTCAGGCCAGGTATTTTATAGGAGAT 4196

4089 ATTATTAATTTCTGGTTGGTATATTCAGCCCACTTCAGGCCAGATATTTTGGAGAT 4148  
4197 TATTTCTTCAGAGACAAGTGAAGATGATGATTCATCTGAAGTGTATGGTTACTGTCCTCAG 4256  
4149 TATTTCTTCAGAGTCAAGTGAAGATGATGATTCATCTGAAGTGTATGGTTACTGTCCTCAG 4208  
4257 ATAAACCCCTTTGGCCAGATCTACATTCAGGAAACATTTTGAATTTTGAAGTGTGTC 4316  
4209 ATAAACCCCTTTGGCCAGATCTACATTCAGGAAACATTTTGAATTTTGAAGTGTGTC 4268  
4317 AAAGGAATGAGTCAAGTGAAGATGATGATTCATCTGAAGTGTATGGTTACTGTCCTCAG 4376  
4269 AAAGGAATGAGTCAAGTGAAGATGATGATTCATCTGAAGTGTATGGTTACTGTCCTCAG 4328  
4377 TTAAGAAACATCTTCAGAGACTGTAAAGAACTACCTGCGAGAACTCAAACGAAAGTTG 4436  
4329 TTAAGAAACATCTTCAGAGACTGTAAAGAACTACCTGCGAGAACTCAAACGAAAGTTG 4388  
4437 TGTTTTGTCTAAGTATCTAGGATCTCTCAGATTTCTGCTAGATGACCACTTACA 4496  
4389 TGTTTTGTCTAAGTATCTAGGATCTCTCAGATTTCTGCTAGATGACCACTTACA 4448  
4497 GGTATGGATCCCAAGCCCAACAGACATGTCGGGAGCAATTCGATTCGATTTAAAC 4556  
4449 GGTATGGATCCCAAGCCCAACAGACATGTCGGGAGCAATTCGATTCGATTTAAAC 4508  
4557 AGAAGCGGGCTGCTATTCGACACTCATATATGAGGAGGAGGAGGCTGCTGTGAT 4616  
4509 AGAAGCGGGCTGCTATTCGACACTCATATATGAGGAGGAGGAGGCTGCTGTGAT 4568  
4617 CGAGTAGCTATCATGCTGCTGGGAGTAAAGATGATCGGAACTGACACATCTAAAG 4676  
4569 CGAGTAGCTATCATGCTGCTGGGAGTAAAGATGATCGGAACTGACACATCTAAAG 4628  
4677 AGTAAATTTGAAAGGCTACTTTTGGAAATTTAAATTTGAAGAGCTGATAGAAAACCTA 4736  
4629 AGTAAATTTGAAAGGCTACTTTTGGAAATTTAAATTTGAAGAGCTGATAGAAAACCTA 4688  
4737 GAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGTCAGGAA 4796  
4689 GAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGTCAGGAA 4748  
4797 AGTTTTCTCTATTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCAA 4856  
4749 AGTTTTCTCTATTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCAA 4808  
4857 TCTTTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCAATTTGAAGATATAGCTTTCT 4916  
4809 TCTTTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCAATTTGAAGATATAGCTTTCT 4868  
4917 CAAGCAACATTTGGACAGGTTTTGTAGAACTCACTAAAGAACTAAGAGGAGGAGATAAT 4976  
4869 CAAGCAACATTTGGACAGGTTTTGTAGAACTCACTAAAGAACTAAGAGGAGGAGATAAT 4928  
4977 AGTTGTGAACCTTTAAACAGACACTTTGGTGGAACTGGAACGAAAGATAGATAGTA 5036  
4929 AGTTGTGAACCTTTAAACAGACACTTTGGTGGAACTGGAACGAAAGATAGATAGTA 4988  
5037 TTTTGAATTTGATTTGTTGGTCTGCTTACTGGAACCTTTCTTTTTCACCTTAATTTTA 5096  
4989 TTTTGAATTTGATTTGTTGGTCTGCTTACTGGAACCTTTCTTTTTCACCTTAATTTTA 5048  
5097 ACTTTGGTTTAAAGATTTTATTTGGAATGTTAACTGGGAGCTGGAACCAAG 5144  
5049 ACTTTGGTTTAAAGATTTTATTTGGAATGTTAACTGGGAGCTGGAACCAAG 5096

RESULT 7  
AX417828  
LOCUS  
DEFINITION  
ACCESSION

AX417828  
Sequence 5 from Patent WO0231147.  
AX417828

5262 bp  
DNA  
linear  
PAT 18-JUN-2002

AX417828.1 GI:21522944  
Homo sapiens (human)  
Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Hu, Y. and Nepomichy, B.  
Human transporter proteins and polynucleotides encoding the same  
Patent: WO 0231147-A 5 18-APR-2002;  
LEXICON GENETICS INC (US)  
FEATURES  
Source  
1..5262  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 91.6%; Score 5014.6; DB 6; Length 5262;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 5014; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
97 AGGTTTATTCAGAAAACATGTCCTCAATTTAGGAGGTAGGAGTTTGGAGACAGACA 156  
238 AGGTTTATTCAGAAAACATGTCCTCAATTTAGGAGGTAGGAGTTTGGAGACAGACA 297  
157 GAACACTTCTAGAGAACTTAAATTAATTCAGAACCAAAAAGAGTAGTGTTCAGG 216  
298 GAACACTTCTAGAGAACTTAAATTAATTCAGAACCAAAAAGAGTAGTGTTCAGG 357  
217 AAATCTTTTCCACTATTTTATTTTGGTGAATTAATTAATTCAGCATGATGCATCAA 276  
358 AAATCTTTTCCACTATTTTATTTTGGTGAATTAATTAATTCAGCATGATGCATCAA 417  
277 ATAAGAAATATGAAGAGTGCCTATATAGAACTCAATCCTATGCAAGAGTTTACTCTTT 336  
418 ATAAGAAATATGAAGAGTGCCTATATAGAACTCAATCCTATGCAAGAGTTTACTCTTT 477  
337 CTAATCTAATCTTGGATATCTCCAGTGACTTAAATTTACAGAGCAGCATCATGCAGAAAG 396  
478 CTAATCTAATCTTGGATATCTCCAGTGACTTAAATTTACAGAGCAGCATCATGCAGAAAG 537  
397 TGTCTACTGATCATCTACTGATGTCATTAATTCAGAAAGATATACAAATGAAAGAAA 456  
538 TGTCTACTGATCATCTACTGATGTCATTAATTTACTGAAAGATATACAAATGAAAGAAA 597  
457 TGTAACTCCAGTCTCTTAAGCCGAGCACTTTTGTAGGTGTGTTTCAAAGACTCCA 516  
598 TGTAACTCCAGTCTCTTAAGCCGAGCACTTTTGTAGGTGTGTTTCAAAGACTCCA 657  
517 TGTCTATGAACCTTCTGTTTCTTGATATGANTCCAGTATCTTCTATTTATATGGATT 576  
658 TGTCTATGAACCTTCTGTTTCTTGATATGANTCCAGTATCTTCTATTTATATGGATT 717  
577 CAAGAGCTGGCTGTTCAAAATCATCTGAGGCTGCTCAGTACTGTCCTCAGGTTTCAAG 636  
718 CAAGAGCTGGCTGTTCAAAATCATCTGAGGCTGCTCAGTACTGTCCTCAGGTTTCAAG 777  
637 TTTTCAAGCATCCATAGATGCTGCCATTTATACAGTTGAAGACCAATGTTTCTTTTGA 696  
778 TTTTCAAGCATCCATAGATGCTGCCATTTATACAGTTGAAGACCAATGTTTCTTTTGA 837  
697 AGGAGCTGGAGTCAACTAAAGCTGTTATTTATGGGAGAACTGCTGTTGTAGAAATAGATA 756  
838 AGGAGCTGGAGTCAACTAAAGCTGTTATTTATGGGAGAACTGCTGTTGTAGAAATAGATA 897  
757 CTTTCCCGGAGGAGTAAATTTTATATACCTAGTTTATAGCATTTTTCACCTTTTGATACT 816  
898 CTTTCCCGGAGGAGTAAATTTTATATACCTAGTTTATAGCATTTTTCACCTTTTGATACT 957  
817 TTTTGGCAATTCATATCGTAGCAGAAAAGAAAAGAAAATTTTAAAGATAA 876  
958 TTTTGGCAATTCATATCGTAGCAGAAAAGAAAAGAAAATTTTAAAGATAA 1017

Qy	877	TGGACATTCATGACTGCTCTTTGGCTTTCTCGGTTCTCTATATACAAAGTTTAAATTT	936
Db	1018	TGGACATTCATGACTGCTCTTTGGCTTTCTCGGTTCTCTATATACAAAGTTTAAATTT	1077
Qy	937	TTCTTATGTCCTCTCTTATGACGATCATTTGGACAGCTCTCTTGTATTTTCCCAAGTA	996
Db	1078	TTCTTATGTCCTCTCTTATGACGATCATTTGGACAGCTCTCTTGTATTTTCCCAAGTA	1137
Qy	997	GCAGCATTTGATATTTCTGCTTTTCTTTTATGATTAATCATCTGTATTTTGGCTT	1056
Db	1138	GCAGCATTTGATATTTCTGCTTTTCTTTTATGATTAATCATCTGTATTTTGGCTT	1197
Qy	1057	TAATGCTGACACCTCTTTTAAAAAATCAAAACATATGGAATAGTTGAAATTTTGTTA	1116
Db	1198	TAATGCTGACACCTCTTTTAAAAAATCAAAACATATGGAATAGTTGAAATTTTGTTA	1257
Qy	1117	CTGTGGCTTTTGGATTTATTTGGCCCTTATGATAATCCCTCATAGAAAGTTTCCCAATCGT	1176
Db	1258	CTGTGGCTTTTGGATTTATTTGGCCCTTATGATAATCCCTCATAGAAAGTTTCCCAATCGT	1317
Qy	1177	TAGTGTGCTTTTTCAGTCTCTTTCTGTCACGTACTTTTGTGATTTGATTTGACAGGTCA	1236
Db	1318	TAGTGTGCTTTTTCAGTCTCTTTCTGTCACGTACTTTTGTGATTTGATTTGACAGGTCA	1377
Qy	1237	TGCAATTTAGAAAGATTTTAAAGAGTCTTCAATTTTCAAATTTGACGTGAGGCCATATC	1296
Db	1378	TGCAATTTAGAAAGATTTTAAAGAGTCTTCAATTTTCAAATTTGACGTGAGGCCATATC	1437
Qy	1297	CTCTAAATTTATACAAATTTATCATGCTCACCTTAAATAGTATATCTATGCTCTTGGCTG	1356
Db	1438	CTCTAAATTTATACAAATTTATCATGCTCACCTTAAATAGTATATCTATGCTCTTGGCTG	1497
Qy	1357	TCTATCTTGATCAAGTCAATTTCCAGGGAAATTTGGCTTACCGAGATCATCTTATATTTTC	1416
Db	1498	TCTATCTTGATCAAGTCAATTTCCAGGGAAATTTGGCTTACCGAGATCATCTTATATTTTC	1557
Qy	1417	TGAAGCTTTCATATTTGTTCAAGAGCAAAAGAAATTTATGAGGAGTTATCAGAGGGCAATG	1476
Db	1558	TGAAGCTTTCATATTTGTTCAAGAGCAAAAGAAATTTATGAGGAGTTATCAGAGGGCAATG	1617
Qy	1477	TTAATGGAATPATTTAGTTTATGAAATTTATGAGCAGTTTCTTTCAGAAATTTGTAGGAA	1536
Db	1618	TTAATGGAATPATTTAGTTTATGAAATTTATGAGCAGTTTCTTTCAGAAATTTGTAGGAA	1677
Qy	1537	AAGAGCCATPAAGAAATTTAGTGTATTTCCAGAGCATACAGAAAGAGGGTGAATGTGG	1596
Db	1678	AAGAGCCATPAAGAAATTTAGTGTATTTCCAGAGCATACAGAAAGAGGGTGAATGTGG	1737
Qy	1597	AGGCTTTTGAGAAATTTGTCATTTGACATATATGAGGCTCAGATTAATGCTTACTTTGGCC	1656
Db	1738	AGGCTTTTGAGAAATTTGTCATTTGACATATATGAGGCTCAGATTAATGCTTACTTTGGCC	1797
Qy	1657	ACAGTGAACAGGAAAGATGATTAATGATTAATTTCTTTGTGACCTCTGCCCACCTTCTG	1716
Db	1798	ACAGTGAACAGGAAAGATGATTAATGATTAATTTCTTTGTGACCTCTGCCCACCTTCTG	1857
Qy	1717	ATGGGTTTGCATCTATATATGGAACACAGCTCTCAGAAATAGATGAATTTTGAAGCAA	1776
Db	1858	ATGGGTTTGCATCTATATATGGAACACAGCTCTCAGAAATAGATGAATTTTGAAGCAA	1917
Qy	1777	GAAAAATGATTTGGCATTTTGCACAGTTAGATATACATTTGATTTTGAAGTAGAAG	1836
Db	1918	GAAAAATGATTTGGCATTTTGCACAGTTAGATATACATTTGATTTTGAAGTAGAAG	1977
Qy	1837	AAAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAATATATATACAGAGAG	1896
Db	1978	AAAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAATATATATACAGAGAG	2037
Qy	1897	TGCAGAGGTTTACTAGATTTAGACATGCAGACTATCAAGATACCAAGCTTAAAAAT	1956
Db	2038	TGCAGAGGTTTACTAGATTTAGACATGCAGACTATCAAGATACCAAGCTTAAAAAT	2097

Qy	1957	TAAGTGTGGTCAAAAAAGAAAGCTGTCTATAGGAATGCTGTTCTTGGAAACCCAAAGA	2016
Db	2098	TAAGTGTGGTCAAAAAAGAAAGCTGTCTATAGGAATGCTGTTCTTGGAAACCCAAAGA	2157
Qy	2017	TACTGCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATGTATGGA	2076
Db	2158	TACTGCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATGTATGGA	2217
Qy	2077	ATCTTTTAAAAATACAGAAAAAGCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATG	2136
Db	2218	ATCTTTTAAAAATACAGAAAAAGCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATG	2277
Qy	2137	AAGCTGACATCTTTCGACATAGGAAAGCTGTGATATCAAGGAATGCTGAAATGTGTG	2196
Db	2278	AAGCTGACATCTTTCGACATAGGAAAGCTGTGATATCAAGGAATGCTGAAATGTGTG	2337
Qy	2197	GTCTTTCAATGTTCTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG	2256
Db	2338	GTCTTTCAATGTTCTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG	2397
Qy	2257	ACAAATATTTGTGCCACAGAAATCTCTTTTCACTGGTTTAAACAAATATACCTGGAGCTA	2316
Db	2398	ACAAATATTTGTGCCACAGAAATCTCTTTTCACTGGTTTAAACAAATATACCTGGAGCTA	2457
Qy	2317	CTTTATACACAGAAATGACCAACAACTGTGTATAGCTTGCCTTTCAAGGACATGAGACA	2376
Db	2458	CTTTATACACAGAAATGACCAACAACTGTGTATAGCTTGCCTTTCAAGGACATGAGACA	2517
Qy	2377	AAATTTTCAGGTTGTTTCTGCTCCCTAGACAGTCAATTTCAAAATTTGGGTGTCTTCTTATG	2436
Db	2518	AAATTTTCAGGTTGTTTCTGCTCCCTAGACAGTCAATTTCAAAATTTGGGTGTCTTCTTATG	2577
Qy	2437	GTGTTTCCATGACGACTTTTGAAGACGTATTTTAAAGCTAGAGTTGGAAGCAGAAATG	2496
Db	2578	GTGTTTCCATGACGACTTTTGAAGACGTATTTTAAAGCTAGAGTTGGAAGCAGAAATG	2637
Qy	2497	ACCAAGCAGATTATAGTGTATTTACTCAGCAGCCACTGGAGGAGAAATGGAATCAAAAT	2556
Db	2638	ACCAAGCAGATTATAGTGTATTTACTCAGCAGCCACTGGAGGAGAAATGGAATCAAAAT	2697
Qy	2557	CTTTTGATGAAATGGAAACAGAGCTTACTTATTTCTTTTGAACCAAGGCTTCTCTAGTGA	2616
Db	2698	CTTTTGATGAAATGGAAACAGAGCTTACTTATTTCTTTTGAACCAAGGCTTCTCTAGTGA	2757
Qy	2617	GCACCATCAGCCTTTGGAAACACAGAGTGTATCAATAGCAAGATTTTCAATTTCTTACT	2676
Db	2758	GCACCATCAGCCTTTGGAAACACAGAGTGTATCAATAGCAAGATTTTCAATTTCTTACT	2817
Qy	2677	TGAAACGTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG	2736
Db	2818	TGAAACGTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG	2877
Qy	2737	TTTCAGATTTTATGTTTGTGTTTCTATCACTCTTTTAAAAATGCTGTGCTCCCATCAAC	2796
Db	2878	TTTCAGATTTTATGTTTGTGTTTCTATCACTCTTTTAAAAATGCTGTGCTCCCATCAAC	2937
Qy	2797	TGTTTCCAGACTTATTTTCTTAAAACTGGAGACAAACCAATAAATAACAAACAGTC	2856
Db	2938	TGTTTCCAGACTTATTTTCTTAAAACTGGAGACAAACCAATAAATAACAAACAGTC	2997
Qy	2857	TGCTTCTTCAAAATTTCTGCTGACTCAGATATCAGTGATCTTATTTAGCTTTTTCAGGCC	2916
Db	2998	TGCTTCTTCAAAATTTCTGCTGACTCAGATATCAGTGATCTTATTTAGCTTTTTCAGGCC	3057
Qy	2917	AGAACATATGCTGACGATGATTAATGACAGTGTACTGTATCCGTGGCTCCCATAGTG	2976
Db	3058	AGAACATATGCTGACGATGATTAATGACAGTGTACTGTATCCGTGGCTCCCATAGTG	3117
Qy	2977	CGGCTTTAAATGTGATGATTCAGAAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTA	3036
Db	3118	CGGCTTTAAATGTGATGATTCAGAAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTA	3177
Qy	3037	CTATGGTTTATTTCTTTTACCATATTTAGTGAATATCATTTAGTAACTACTATCTTTTATCAT	3096



KEYWORDS	Homo sapiens (human)	953
SOURCE	Homo sapiens	840
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 Hu, Y. and Nepomnichy, B.	
AUTHORS	Human transporter proteins and polynucleotides encoding the same	
TITLE	Patent: WO 0231147-A 1 18-APR-2002;	
JOURNAL	LEXICON GENETICS INC (US)	
FEATURES	Location/Qualifiers	
source	1. .4929	
ORIGIN	/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"	
Query Match	89.9%; Score 4922.6; DB 6; Length 4929;	
Best Local Similarity	99.9%; Pred. No. 0;	
Matches 4922; Conservative	4; Mismatches 3; Indels 0; Gaps 0;	
QY	114 ATGTCCACTGCAATTAGGAGAGTAGAGTTGGAGACACAGACGACCACTTCTACTGAAG	173
DB	1 ATGTCCACTGCAATTAGGAGAGTAGAGTTGGAGACACAGACGACCACTTCTACTGAAG	60
QY	174 AATTACTTAATTAATGACAGACCAAAAGAGTAGTGTTCAGGAAATCTTTTTCCACTA	233
DB	61 AATTACTTAATTAATGACAGACCAAAAGAGTAGTGTTCAGGAAATCTTTTTCCACTA	120
QY	234 TTTTCTTTTATTTGGTTAATATTAATTAGCATGATGCATCCAAATAGAAATATGAAGAA	293
DB	121 TTTTCTTTTATTTGGTTAATATTAATTAGCATGATGCATCCAAATAGAAATATGAAGAA	180
QY	294 GTCCCTAATATAGAACTCAATCCTATGGACAGTTTACTCTTTCTAATCTAATCTTTGGA	353
DB	181 GTCCCTAATATAGAACTCAATCCTATGGACAGTTTACTCTTTCTAATCTAATCTTTGGA	240
QY	354 TATACTCCAGTGACTAATATTAACAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA	413
DB	241 TATACTCCAGTGACTAATATTAACAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA	300
QY	414 CCTGATGTCAATTAATTAACAGAAATATACAAATGAAAGAAATGTTAAATCCAGTCTC	473
DB	301 CCTGATGTCAATTAATTAACAGAAATATACAAATGAAAGAAATGTTAAATCCAGTCTC	360
QY	474 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAGAAAGTCCATGCTCCTATGAACITTCGT	533
DB	361 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAGAAAGTCCATGCTCCTATGAACITTCGT	420
QY	534 TTTTCTTCCATGATAGTCCAGTATCTCTAATATATATGGAATTCAGAGCTGGCTGTTC	593
DB	421 TTTTCTTCCATGATAGTCCAGTATCTCTAATATATATGGAATTCAGAGCTGGCTGTTC	480
QY	594 AAATCATGTGAGCTGCTCAGTACTGCTCCTCAGTGTTCACAGTGTTCAGAGCAATCCATA	653
DB	481 AAATCATGTGAGCTGCTCAGTACTGCTCCTCAGTGTTCACAGTGTTCAGAGCAATCCATA	540
QY	654 GATGCTGCCATTTACAGTTGAAGACCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACT	713
DB	541 GATGCTGCCATTTACAGTTGAAGACCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACT	600
QY	714 AAGCTGTTATATGGGAAACTGCTGCTGTAGAAATAGATACCTTTCCCGAGGAGTA	773
DB	601 AAGCTGTTATATGGGAAACTGCTGCTGTAGAAATAGATACCTTTCCCGAGGAGTA	660
QY	774 ATTTTAATATACCTAGTTATAGCATTTTCCACTTTTGGATACCTTTTGGCAATTCATATC	833
DB	661 ATTTTAATATACCTAGTTATAGCATTTTCCACTTTTGGATACCTTTTGGCAATTCATATC	720
QY	834 GTAGCAGAAAGAAAGAAATTAAGAAATTTTAAAGATAATGGGACTTCATGATACT	893
DB	721 GTAGCAGAAAGAAAGAAATTAAGAAATTTTAAAGATAATGGGACTTCATGATACT	780

Db 1861 AGAAGCTGTCATTAGGAATGCTGTTCTTGGGAACCCAAAGATACCTGCTGATGAA 1920  
Qy 2034 CCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAAAACAGA 2093  
Db 1921 CCAACAGCTGGAATGGAACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAAAACAGA 1980  
Qy 2094 AAAGCCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGCATTTCTTGCA 2153  
Db 1981 AAAGCCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGCATTTCTTGCA 2040  
Qy 2154 GATAGAAAGCTGTGATATCAACAGGAATGCTGAAATCTGTGGTCTTCCAAATGTTCTC 2213  
Db 2041 GATAGAAAGCTGTGATATCAACAGGAATGCTGAAATGCTGTGGTCTTCCAAATGTTCTC 2100  
Qy 2214 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACA 2273  
Db 2101 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACA 2160  
Qy 2274 GAATCTCTTTCTCACTGGTTAAACAAACATATACCTGGAGCTACTTTATTAACAAGAA 2333  
Db 2161 GAATCTCTTTCTCACTGGTTAAACAAACATATACCTGGAGCTACTTTATTAACAAGAA 2220  
Qy 2334 GACCAACACTGCTGTATAGCTTGCCTTCAAGGACATGGACAAATTTCAAGTTTGT 2393  
Db 2221 GACCAACACTGCTGTATAGCTTGCCTTCAAGGACATGGACAAATTTCAAGTTTGT 2280  
Qy 2394 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGTGTTCACACGACT 2453  
Db 2281 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTTATGGGTGTTCACACGACT 2340  
Qy 2454 TTGGAAGAGCTATTTTAAAGCTAGAAGTTGAGCAGAAATTTGACCAAGCAGATTTAGT 2513  
Db 2341 TTGGAAGAGCTATTTTAAAGCTAGAAGTTGAGCAGAAATTTGACCAAGCAGATTTAGT 2400  
Qy 2514 GTATTACTCAGCAGCCTGAGGAAGAAATGGATTCAAAATCTTTTGTGAAATGGAA 2573  
Db 2401 GTATTACTCAGCAGCCTGAGGAAGAAATGGATTCAAAATCTTTTGTGAAATGGAA 2460  
Qy 2574 CAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTAGTGAGCACCATGAGCCTTTGG 2633  
Db 2461 CAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTAGTGAGCACCATGAGCCTTTGG 2520  
Qy 2634 AAACACAGATGTATACATACCAAGTTTCATTTCTTACCTTGAAACGTTGAAATGAA 2693  
Db 2521 AAACACAGATGTATACATACCAAGTTTCATTTCTTTTACCTTGAAACGTTGAAATGAA 2580  
Qy 2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2753  
Db 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2640  
Qy 2754 TTGGTTTCATCTCTTTTAAATAATGCTGGTTCCTCCATCAAACTGTTTCCAGACTTAT 2813  
Db 2641 TTGGTTTCATCTCTTTTAAATAATGCTGGTTCCTCCATCAAACTGTTTCCAGACTTAT 2700  
Qy 2814 TTTCTAAAAACCTGGAGACAAACCATAAATACAAACAAAGCTGCTGTTCTTCAAAATCT 2873  
Db 2701 TTTCTAAAAACCTGGAGACAAACCATAAATACAAACAAAGCTGCTGTTCTTCAAAATCT 2760  
Qy 2874 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAGCCAGAACATTAATGGTGAC 2933  
Db 2761 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAGCCAGAACATTAATGGTGAC 2820  
Qy 2934 ATGATTAAATGACAGTGACTATGTATCCGTTGGTTCCTCCATAGTGGCTTTTAAATGTGAT 2993  
Db 2821 ATGATTAAATGACAGTGACTATGTATCCGTTGGTTCCTCCATAGTGGCTTTTAAATGTGAT 2880  
Qy 2994 CATTCAGAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTATGTTTATCTTTA 3053  
Db 2881 CATTCAGAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTATGTTTATCTTTA 2940  
Qy 3054 CCTATATTAGTGAATATCAATTAGTAACCTATCTTTATCATTTTAAATGTGACTGAAACC 3113

Db 2941 CCTATTATTAGTGAATATCATATTAGTAACCTACTATCTTTATCATTTAAATGTGACTGAAACC 3000  
Qy 3114 ATCCAGATCTGAGATGATACCCCATCTTTCAAGAAATTAAGTATATAGTTTTTAAAAATGAG 3173  
Db 3001 ATCCAGATCTGAGATGATACCCCATCTTTCAAGAAATTAAGTATATAGTTTTTAAAAATGAG 3060  
Qy 3174 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTTGCATGCAATGCCACTTACTTTGCC 3233  
Db 3061 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTTGCATGCAATGCCACTTACTTTGCC 3120  
Qy 3234 ATGGAATAATGACAGAGAAATCATTAAGATCAAAAGCTTATATCTCAACTTAAACTTTTCAAGTCT 3293  
Db 3121 ATGGAATAATGACAGAGAAATCATTAAGATCAAAAGCTTATATCTCAACTTAAACTTTTCAAGTCT 3180  
Qy 3294 TTGCCATCTGCAATTTGGAAATGGAACAAGCTGTTGTTGATATCCCTTATTTTTTATCAT 3353  
Db 3181 TTGCCATCTGCAATTTGGAAATGGAACAAGCTGTTGTTGATATCCCTTATTTTTTATCAT 3240  
Qy 3354 CTTATTTTGAATGCTAGGAAGCTTATTTGGCAATTCATATGGAATTAATTTTTTATCTGTA 3413  
Db 3241 CTTATTTTGAATGCTAGGAAGCTTATTTGGCAATTTCTATGGAATTAATTTTTTATCTGTA 3300  
Qy 3414 AGTTTCCTTGTGCTGTTTTTTTGGCTTATTTGGTATGTTTCCATCAGTATTTCTGTTCACT 3473  
Db 3301 AGTTTCCTTGTGCTGTTTTTTTGGCTTATTTGGTATGTTTCCATCAGTATTTCTGTTCACT 3360  
Qy 3474 TATATTGCTTCTTTCACTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3533  
Db 3361 TATATTGCTTCTTTCACTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
Qy 3534 TATTTCTGCGCAGCTGTTGCTTGTATGCAATCACTGAAATTAACCTTTCTTTATGGATAC 3593  
Db 3421 TATTTCTGCGCAGCTGTTGCTTGTATGCAATCACTGAAATTAACCTTTCTTTATGGATAC 3480  
Qy 3594 ACAATTGCAACTATTTCTTCAATTTGCTATGCTATCATCATTTCCAACTATCCACTTCTA 3653  
Db 3481 ACAATTGCAACTATTTCTTCAATTTGCTATGCTATCATCATTTCCAACTATCCACTTCTA 3540  
Qy 3654 GGTGCTGCTGATTTCTTCAATAAGATTTCTTGGAGAAATGTAGCAAAAATGTGGACACC 3713  
Db 3541 GGTGCTGCTGATTTCTTCAATAAGATTTCTTGGAGAAATGTAGCAAAAATGTGGACACC 3600  
Qy 3714 TATAATCCATGGGATAGGCTTTTCACTAGTGTATATTCGCTTACCTGAGTGTGTACTG 3773  
Db 3601 TATAATCCATGGGATAGGCTTTTCACTAGTGTATATTCGCTTACCTGAGTGTGTACTG 3660  
Qy 3774 TGGATTTTCTCTTCAATACTATGAAAAAATATGGAGGAGATCAATAGAAAGAT 3833  
Db 3661 TGGATTTTCTCTTCAATACTATGAAAAAATATGGAGGAGATCAATAGAAAGAT 3720  
Qy 3834 CCCTTTTTCAGAAACCTTTTCAACGAGTCTAAAAATAGGAAGCTTCCAGAACCCAGAC 3893  
Db 3721 CCCTTTTTCAGAAACCTTTTCAACGAGTCTAAAAATAGGAAGCTTCCAGAACCCAGAC 3780  
Qy 3894 AATGAGATGAAGATGAAGATGTAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 3953  
Db 3781 AATGAGATGAAGATGAAGATGTAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCTTAAAGCT 3840  
Qy 3954 TGCAGTGTGTGAGAGAAACCATCATTTATGCTGAGCAATTTGCTATGAAAGATATGAT 4013  
Db 3841 TGCAGTGTGTGAGAGAAACCATCATTTATGCTGAGCAATTTGCTATGAAAGATATGAT 3900  
Qy 4014 GACAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTAAAATACATCTCT 4073  
Db 3901 GACAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCAACTAAAATACATCTCT 3960  
Qy 4074 TTCTGTGTGAAAAAGAGAGATCTTTAGGACTATTTGGGTCCAAATGGTGTCTGGCAAAAGC 4133  
Db 3961 TTCTGTGTGAAAAAGAGAGATCTTTAGGACTATTTGGGTCCAAATGGTGTCTGGCAAAAGC 4020  
Qy 4134 ACAATTATTAAATTTCTGTTGTTGATATTTGAACCAACTTTCAGGCGAGGATTTTTTAGGA 4193  
Db 4021 ACAATTATTAAATTTCTGTTGTTGATATTTGAACCAACTTTCAGGCGAGGATTTTTTAGGA 4080











FEATURES	source	Location/Qualifiers
	1. .4785	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
ORIGIN		
Query Match	86.9%;	Score 4759; DB 6; Length 4785;
Best Local Similarity	99.8%;	Pred. No. 0;
Matches 4759; Conservative	4;	Mismatches 4; Indels 0; Gaps 0;
QY	114	ATGTCCTACTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACCAGAACACTTCTACTCTGAAG 173
Db	1	ATGTCCTACTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACCAGAACACTTCTACTCTGAAG 60
QY	174	AATTACTTTAATTAATCGAGAACCAAAAAGAGTAGTGTTCAGGAAATTCCTTTTCCACTA 233
Db	61	AATTACTTTAATTAATCGAGAACCAAAAAGAGTAGTGTTCAGGAAATTCCTTTTCCACTA 120
QY	234	TTTTTTTTTATTTTGGTTTAATTAATTAGCATGATGCATCCAAATAGGAATATGAGAA 293
Db	121	TTTTTTTTTATTTTGGTTTAATTAATTAGCATGATGCATCCAAATAGGAATATGAGAA 180
QY	294	GTGCCTAATATAGAACTCAATCCTATGGACAAGTTTACTCTTTCTAATCTAATCTTTGGGA 353
Db	181	GTGCCTAATATAGAACTCAATCCTATGGACAAGTTTACTCTTTCTAATCTAATCTTTGGGA 240
QY	354	TATACTCCAGTGACTAATATATACAAAGCAGCATCATGCAGAAAGTGCTCTACTGATCATCTA 413
Db	241	TATACTCCAGTGACTAATATATACAAAGCAGCATCATGCAGAAAGTGCTCTACTGATCATCTA 300
QY	414	CCTGATGTCTAATAATTCTGGAAGATATACAAATCAAAAAGAAATGTTTAAACATCCAGTCTC 473
Db	301	CCTGATGTCTAATAATTCTGGAAGATATACAAATCAAAAAGAAATGTTTAAACATCCAGTCTC 360
QY	474	TCTAAGCGGAGCAACTTTGTAGGTGTGGTTTTCAAAGCTCCATGTCTCTATGAATCTTCGT 533
Db	361	TCTAAGCGGAGCAACTTTGTAGGTGTGGTTTTCAAAGCTCCATGTCTCTATGAATCTTCGT 420
QY	534	TTTTTTTCTGTATGATTCAGTATCTTCATTTTATATGGATTCAAGAGCTGGCGTGTTC 593
Db	421	TTTTTTTCTGTATGATTCAGTATCTTCATTTTATATGGATTCAAGAGCTGGCGTGTTC 480
QY	594	AAATCATGTGAGGCTGCTCAGTACTGCTCAGTTCCTCAGTTTTCACAGTTTTCACAGCATCCATA 653
Db	481	AAATCATGTGAGGCTGCTCAGTACTGCTCAGTTCCTCAGTTCCTCAGTTCCTCAGCATCCATA 540
QY	654	GATGCTGCCATTATACAGTTTGAAACCAATGTTTCTCTTTGGAGAGGCTGGAGTCAACT 713
Db	541	GATGCTGCCATTATACAGTTTGAAACCAATGTTTCTCTTTGGAGAGGCTGGAGTCAACT 600
QY	714	AAAGCTGTATTATTAGGAGAAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGGAGTA 773
Db	601	AAAGCTGTATTATTAGGAGAAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGGAGTA 660
QY	774	ATTTTAATATACCTTAGTTATAGCAATTTTACCTTTTGGATACCTTTTGGCAATTCATATC 833
Db	661	ATTTTAATATACCTTAGTTATAGCAATTTTACCTTTTGGATACCTTTTGGCAATTCATATC 720
QY	834	GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATATGGCACTTCATGATACT 893
Db	721	GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATATGGCACTTCATGATACT 780
QY	894	GCCTTTTGGCTTTCCTGGTTTCTCTATATATACAAAGTTTAAATTTTCTTATGTCCCTTCTT 953
Db	781	GCCTTTTGGCTTTCCTGGTTTCTCTATATATACAAAGTTTAAATTTTCTTATGTCCCTTCTT 840
QY	954	ATGCGAGTTCATTGGCAGAGCTTCCTTTGTTATTTCTCCTCAAGCTAGCAGCATTTGTGATATTT 1013
Db	841	ATGCGAGTTCATTGGCAGAGCTTCCTTTGTTATTTCTCCTCAAGCTAGCAGCATTTGTGATATTT 900
QY	1014	CTGCTTTTTTTTCTTTATGGAATATCATCTGTATTTTTTGTCTTTTAAATGCTGACACCTCTT 1073





QY 587 CTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTCCTCAGGTTTTCACAGTTTTCACAAGC 646  
DB 683 CTGTTCAAGACATGTGATGTGCTCAGTACTGTCCTTGGGGTTTACAGTTCTGTCAGGC 742  
QY 647 ATCCATAGATGCTGCCATTTATACAGTCTGAAGACCAATGTTTCTCTTTGGAAGGAGCTGGA 706  
DB 743 ATCGATAGATGCTGCCATTTATACAGTCTGGAGCCAAATGTTTCTGTGTGGAGGAGCTGGA 802  
QY 707 GTCAATAGATGCTGTTATTTATGGGAGAACTGCTGTTGTAGAAATAGATATAGATATCTTCCCG 766  
DB 803 GTGCAACCAAGCTGTGATCATGGAGAGGCGGCTGTTGTGGAGATTTGACACCTTCCCGCG 862  
QY 767 AGGAGTAAATTTAATATACCTAGTTATAGCAATTTTCACTTTTTCGATACTTTTGGCAAT 826  
DB 863 AGGGGTATCTCTACCTCTGCTCATAGCCCTTCTGCGCCCTTGGCTACTTCTCTGGCAAT 922  
QY 827 TCATATCTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAATAGGACTTCA 886  
DB 923 CCACATCTGGCAGAAAAAGAAAGAGTTAAAGGAATTTTAAAGATAATAGGACTTCA 982  
QY 887 TGATATCTGCTTTTGGCTTCTGCTGTTCTCTATATACAGTTTAAATTTTCTTATGTC 946  
DB 983 TGACACTGCTTTTGGCTTCTGCTGTTCTCTGACGCAAGCTTGATTTTCTTATGTC 1042  
QY 947 CTTCTTATGGCAGTCAATGGACAGCTTCTTTGTTATTTCTCCTCAAAGTAGCAGCATTTG 1006  
DB 1043 CTTGCTTATGGCTGTCATCGCAACAGCTTCTCGTTATTTCTCCTCAGAGTAGCAGCATTTG 1102  
QY 1007 GATATTTCTGCTTTTCTTATAGGATATCATCTGTATTTTGGTTTAAATGCTGAC 1066  
DB 1103 GATCTTTCTACTGGCTTCTTATATGGATTTGTCATCTGTGTTTTTGGCTTTAAATGCTGAC 1162  
QY 1067 ACCTCTTTTAAATAATCAAAACATGTGGGAATAGTTGAAATTTTGTACTGTGGCTTT 1126  
DB 1163 GCCTCTTTTAAATAATCAAAACATGTGGAGTGGTTGAGTTTGTTCACCGTGGTT 1222  
QY 1127 TGGATTTATGGCTTATGATATCCTCATAGAAAGTTTCCCAAATGGTTAGTGGCT 1186  
DB 1223 TGGATTTATGGCTGCTGATGTCCTCATAGAAAGTTTCCCAAAGTGGCTGGTGGCT 1282  
QY 1187 TTTCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1246  
DB 1283 CTTCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1342  
QY 1247 AGATTTTAAATGAGTGTCTCAATTTCAAATTTGATGCTGAGGCTGCTGCTGCTGCTGCT 1306  
DB 1343 AGATTTTAAATGAGTGTCTCAATTTCAAATTTGATGCTGAGGCTGCTGCTGCTGCTGCT 1402  
QY 1307 TACAATTTATGCTCAGCTTATAGTATATTTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1366  
DB 1403 TACTATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462  
QY 1367 TCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAAGCTTTC 1426  
DB 1463 CCAAGTCAATTCAGGGGAATTTGGCTTACGGAGATCATCTTTGATTTTCTGAAAGCTTTC 1522  
QY 1427 ATATTTGCTCAAGAGCAAAAGAAATTTAGAGGATTTACAGAGGCAATGTTATGGA 1486  
DB 1523 GTATTTGGTCAAAAAAATAAAGAAATTTATAGAGGCTATCGAGGGCAACATTTATGGA 1582  
QY 1487 TATTAGTTTATGTAATTTATGAGCTTCTTCAAGTTTGTAGGAAAGAGGCAAT 1546  
DB 1583 TATTAGTCTCAATGAAATTTGTAGCCGCTTCTTCAAGTTTGTAGGAAAGAGGCTAT 1642  
QY 1547 AAGAAATTTAGTGTATTTAGAGACATACAGAAAGAGGCTGAAATGTTGAGGCTTTGAG 1606  
DB 1643 AAGAAATTTAGTGTATTTAGAGATTTATAGAGGCTTCTGCTTCTTCTGAGGCTTTGAG 1702  
QY 1607 AATTTGCTATTTGATATTTAGAGGCTGAGTTTCTGCTTCTTCTGAGGCTTTGAG 1666  
DB 1703 AATTTGCTATTTGATATTTAGAGGCTGAGTTTCTGCTTCTTCTGAGGCTTTGAG 1762  
QY 1667 AGGAAAGAGTACATTTGATGATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTGC 1726

DB 1763 AGGAAAGAGTACATTTGATGATATTTCTGTGTGGACCGTCTCCACCTCTGATGGGTTGC 1822  
QY 1727 ATCTATATATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGAT 1786  
DB 1823 TTCTATATATGACACAGAGTCTCTGAAATAGATGAAATATTTGAAGCAAGAAAAATGAT 1882  
QY 1787 TGGCATTTGTCACAGTATAGATATACATTTGATGTTTGTGACAGTATAGAAATTTATC 1846  
DB 1883 TGGCATATTTGTCGAGTATAGATATTAACCTTTGATGTTTGTGACAGTATAGAAATTTATC 1942  
QY 1847 AATTTGGCTTCAATCAAGAGGATACAGCCCAACAAATAAATCAAGAAAGTGCAGAGGT 1906  
DB 1943 AATTTGGCTTCAATCAAGAGGATACAGCCCAACAAATAAATCAAGAAAGTGCAGAGGT 2002  
QY 1907 TTTATAGATTTTATAGATGACATATCAAGATTAACCAAGCTTAAATAATTAAGTGGTGG 1966  
DB 2003 GTTGTCTGGATCTGGACATCAAGCCATCAAGATTAATCAAGCGAAATTTAAGCGGTGG 2062  
QY 1967 TCAAAAAAGAAAGCTCTCATTAGGAATTCCTTCTGGGAACCCAAAGATATCTGCTGCT 2026  
DB 2063 TCAGAAAAAGAGTCTCTGTAGGAATTCAGTCTCTGGGAATTCAGAGTATCTCTGCT 2122  
QY 2027 AGATGAACCAACAGCTGGAAATGGAACCTCTGTTCTCGACATATTTGTATGAAATCTTTTAA 2086  
DB 2123 AGACGAGCTCAGCAGGAATGGAACCTCTGCTCTGCAATATTTGTTGAAATCTTTCTAAA 2182  
QY 2087 ATACAGAAAGCCATTCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAAGTGCAT 2146  
DB 2183 GTATGAAGAGGCTTACAGAGTACCGGTGTTTAGTACTCATCTTCATGGATGAGGCTGACAT 2242  
QY 2147 TCTTGAGATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTTGTTGTTCTTCAAT 2206  
DB 2243 TCTTGCGCAGAGGAAGCTGTCTATATCAAGGAATGCTGAAAGTGTGTTGTTCTTCAAT 2302  
QY 2207 GTTCTCTCAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATG 2266  
DB 2303 TTTCTCTAAAGAGTAAATGGGGAATCGGCTACCGCTGAGCATGTATATAGACAGATCTG 2362  
QY 2267 TGCCACAGAAATCTCTTTCTCACTGGTTAAAACAAATATACCTGAGGCTTCTTTATACA 2326  
DB 2363 TGCCACAGAGTCACTGTCTGCTGCTGTTAGGAGCAGACATACCCGCGAGCGCTACTGCA 2422  
QY 2327 ACAGAAATGCCAAACAACTTTGTATAGTGTGCTTTCAAGGACATGGACAAATTTTCAGG 2386  
DB 2423 GCAGAAATGCCAGCAGCTGCTGACAGCTGCTTTCAAAGACATGGACAAATTTTCAGG 2482  
QY 2387 TTTGTTTTCTGCTTACAGAGTCAATTTCAAATTTGGGTGTCATTTCTTATGTTGTTTCCAT 2446  
DB 2483 CTTGTTTTCTGCTTACAGATTCATTTCAAACCTGGGTGTTATTTCTTATGTTGTTTCCAT 2542  
QY 2447 GACGACTTTGGAGAGCTATTTTAAAGCTAAGATGAAAGTGAAGCAGAAATTTGACCAAGCAGA 2506  
DB 2543 GACAAATTTGGAGATGTTATTTCTTAAAGCTAGAAGTTGAAGCAGAAATTTGACCAAGCAGA 2602  
QY 2507 TTTATGTTGTTTACTCAGCAGCTCTGGAGGAAGAAATGGAATTTCAAAATCTTTTGAAGA 2566  
DB 2603 CTATGTTGTTTACTCAGCAGCTGCTGAGGAGGAGGAAACAGATTTCAAATCTTTGATGA 2662  
QY 2567 AATGGAACAGAGCTTACTTTTCTGAAAACCAAGGCTTCTTAGTGAAGCAGCAATGAG 2626  
DB 2663 AATGGAACAGAGTTTACTTTTCTCTCTGAAAACCAAGGCTTCTTAGTGAAGCAGCAATGAG 2722  
QY 2627 CTTTGGAAACACAGATGTATACATGCAAGTTTCTTCTTACCTTGAAGCTGA 2686  
DB 2723 CCCTTGGAAACAGAGTGTCTACGATTTGCAAGTTTCTTCTCTCTGAAACCGAGA 2782  
QY 2687 AAGTAAATCAGTGAAGTGTGTTGCTCTGCTTTTAAATTTTTTTTTCACAGTTTCAATTTT 2746  
DB 2783 AAGTAAATCAGTGAAGTGTGTTGCTCTGCTTTTAAATTTTTTTTTCACAGTTTCAATTTT 2842  
QY 2747 TATGTTTTTGGTTTCACTCTTTTAAATTTGCTGTTGTTTCCATCAAACTTTGTTCCAGA 2806



Db 2843 TATGGTTTTCTCCATCTTTTAAAGTGTGGTCCCATCAAACTCGTTCAG 2902  
QY 2807 CTTATATTTCTAAACCTGGAGACAACCAATAAATACAAAAGTCTGCTTCTCA 2866  
Db 2903 CTTGATTTCTAAAGCCTGGAGATAAACCTCATAAATACAAAAGCCTGCTTCA 2962  
QY 2867 AAATCTGCTGACTCAGATATCAGTATCTTATTAGCTTTTTTCAAGCCAGAACATAAT 2936  
Db 2963 AAATCTGCTGACTCAGATATCAATGGTCTTATTAGTTTTTGCACACGACATAT 3022  
QY 2927 GGTGAGATGATTAATGACAGTGAATATGATTCGTGGCTGCCCATAGTGGGCTTAAA 2986  
Db 3023 GGTGGCAATGTTTAAATGACAGTGAATATGATTCGTGGCTGCCCATAGTGGGCTTCAA 3082  
QY 2987 TGTGATGCAATCAGAAAAGGACTATGTTTTTGCAGCTGTTTTCAACAGTACTATGGTTTA 3046  
Db 3083 TGTGGTGGCTTCTGAAAAGGACTATGTTTTTCTGCTGTTTTCAACAGTACTATGGTTTA 3142  
QY 3047 TTCTTTAOCCTATATTAGTGAATPATCATTAGTAACCTACTATCTTTATCATTTAAATGAC 3106  
Db 3143 TTGTTTGCCAGTCATGATGAACATCATTAGTAACCTACTATCTTTATCATTTAAATGAC 3202  
QY 3107 TGAACCACTCCAGATCTGGATGACCCCATCTTTCAAGAAATTAAGTATAGTTTTAA 3166  
Db 3203 TGAAGCCATCCAGACTGGAGTACCCCGTTCAATCAAGAAATTAAGTATAGTTTTAA 3262  
QY 3167 AATTGAGCTGTTTTCAAGCAGCTTTGCTTGGAAATCATTTGTAATGCAATGCCACCTTA 3226  
Db 3263 AATTGAGCTATATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTAATGCAATGCCACCTTA 3322  
QY 3227 CTTTGGCATGGAATGACAGAAATCATAGATCAAGCTTATATCTCAACTTAACTTTC 3286  
Db 3323 CTTTGGCATGGAATGACAGAAATCATAGATCAAGCTTATATCTCAACTTAACTTTC 3382  
QY 3287 AGGTCTTTTGCCATCTGCATATGGATTTGGATGGAAGCTGTTGATATCCCTTATTTTT 3346  
Db 3383 GGGTCTTTTGCCATCAGCCTATGGGTTGGACAAGCTGTTGGATATTCCTTTGTTTT 3442  
QY 3347 TATCATCTTATTTGATGCTAGGAAGCTTATGGCANTTTCATAGGAATATATTTTTA 3406  
Db 3443 TGTGTTCTGATTTGATGCTGGAGTTTATTTGTCATTTATCATGAGCTGATTTTTTA 3502  
QY 3407 TACTGTAAGTTTCTGCTGCTGTTTTTGGCTTATGTTATGTTTCCATCAGTTATCT 3466  
Db 3503 TCTGTCNAATTTCTGCTGCTGTTTTGCTCTATGCTTATGTCCTCGCTCACTCT 3562  
QY 3467 GTTCACTTATATGCTTCTTCCCTTTAAGAAAATTTTAAATACCAAGAAATTTGGTC 3526  
Db 3563 GTTCACTATAGCTTCTGCTTCACTTTCAAGAAAATTTTAAATACCAAGAAATTTGGTC 3622  
QY 3527 ATTTATCTATCTGTCGAGCGTTGGCTTGTATTCGAATCACTGAAATAACTTTCTTTAT 3586  
Db 3623 ATTTATCTATCTGTCGCGAATGGCTTGTGTCANTCATCGGAACACTTTCTTTCT 3682  
QY 3587 GGGATACAAATGCAACTTATCTTCAATPATGCCCTTTTGTATCATPATCCAAATCTATCC 3646  
Db 3683 GCAATATGCAATGAGCTGCTTTCAATATACACCTTCTGATAGCCATTCCAATCTACCC 3742  
QY 3647 ACTTCTAGCTTGCCTGATTTCTTCAAGAAATTTCTTGGAGATGTACGAAATATGT 3706  
Db 3743 TCTCTGGGTGCTGCTGATTTCTTCAAGAGGTTCTTGGAGAAATATGCCAAAATGA 3802  
QY 3707 GGACACTTATATCCATGGATAGGCTTTCACTAGCTGTTATATCGCCTTACCTGCACTG 3766  
Db 3803 GAATACCTAATCCCTGGATAGACTTTTATGCTGTATCATGCTCCCTACCTGCACTG 3862  
QY 3767 TGTACTGTGATTTTCTCTTACATATCATGAGAAAATATGGAGCAGATCAATAG 3826  
Db 3863 TATACTGTGGATTTTCTCTTCAACACTATGAGAAAATACATGGAGCAGATCTATAAG 3922  
QY 3827 AAAGATCCCTTTTTCAGAAAACCTTTCAACGAAGTCTTAAATAGGAAGCTTTCAGAAC 3886  
Db 3923 AAAGATCCATTTTTCAGGGCCCTTTTCAAAAGGCCAAAATAGAAATTTCCAGAAC 3982

QY 3887 ACCAGACATGAGGATGAAGATGTCAAAGCTGAAGACTAAAGCTCAAAGAGCT 3946  
Db 3983 ACCATCATATGAGGATGAAGATGTCAAAGCTGAAGACTGAAGGTTAAAGAGCT 4042  
QY 3947 GATGGGTGCTGAGTGTGTGAGGAGAAAACCATCTTATGTCAGCAATTTTGCATAAAGA 4006  
Db 4043 GATGGGTGCTGAGTGTGTGAGGAGAAAGCCACCATTTATGGTATGCAATTTTGCATAAAGA 4102  
QY 4007 ATATGATGACAGAAGATATTTCTTTTCAAGAAAAGTAAAGAAAGTGGCAACTTAAATA 4066  
Db 4103 GTATGATGACAGAAGATATTTCTTCAATCAAGAAAACCAAAAAGTAGCAACAAAATA 4162  
QY 4067 CATCTCTTCTGTGTGAAAAGGAGAGATCTTAGGACTATTTGGGTCCAAATGGTCTGG 4126  
Db 4163 CATCTCTTCTGTGTGAAAAGGAGAGATCTTAGGACTATTTGGGTCCAAATGGAGCTGG 4222  
QY 4127 CAAAAGCACAATTTAATATTTCTGGTGTGATTTGAAACCAACTTCAGGCCAGGTATT 4186  
Db 4223 CAAAAGCAGATCATTTAACACTCTGGTTGGGACCGTTGAGCCAACTTCAGGCCAGATCTT 4282  
QY 4187 TTTAGGAGATTTCTTTCAGAGACAAGTCAAGATGATGATTTCACTGAAGTGTATGGTTA 4246  
Db 4283 TCTAGGAGATTTATGTTTCACTTCAAGTGAAGATGACGATCCATTTAAGTGTATGGGTA 4342  
QY 4247 CTGCTCTCAGATAAACCCCTTTGTGGCCAGATCTACTCATTTGCGAGAACATTTTGAATTTA 4306  
Db 4343 CTGCCCCAGACAAAACCCACTGTGGCCAGATCTCACTCTACAGGAACATTTTGGATTTA 4402  
QY 4307 TGGAGCTGTCAAGGAATGAGTGAAGTCAAGTCAAGTGAAGAAAGTCAATAGTCGAATTAACA 4366  
Db 4403 CGAGCTGTGAAGGAATGAGTCTGGTGACATGAAGGAAGTCATCAGTTCGAATTAACA 4462  
QY 4367 TGCATTTGATTTTAAAGAACATCTTTCAGAAAGCTGTFAAAGAACTACCTGCGAGGAATCAA 4426  
Db 4463 AGCCCTTCTGTTTGAAGAACATCTTTCAGAAAGCTGTFAAAGAACTACCTGCGAGGAATCAA 4522  
QY 4427 ACCAAGTTGTGTTTTGCTTAAGTATGCTAGGAATCCTCAGATTTACTTTCTAGATCA 4486  
Db 4523 GCAAAAGCTGTGTTTTGCTCTCAGCATGCTGGGAACCCCTCAGGTGACCTGCTGGATGA 4582  
QY 4487 ACCATCTACAGTATGGATCCCAAGCCAAAACAGCAATGTGGCGAGCAATTCGAATGTC 4546  
Db 4583 GCCGTCCACAGGATGACCCCAAGCCAAAGCAGACATGTGGAGAGCTATTCGAATGTC 4642  
QY 4547 ATTTAAAACAGAAAGCGGCTGCTATTTCTGACCCTCACTATATATGGAGAGCGCAGGC 4606  
Db 4643 ATTTAAAACAGAAAGCGGCGCCCTCTGCTGCTACTTCAATCATGGAAGAGCGCAGGC 4702  
QY 4607 TGTCTGTGATCGAGTATGATCATGTTGCTGCGCAGTTAAGATGTATCGGAACAGTACA 4666  
Db 4703 TGTCTGTGACAGAGGGGCCATCATGTAATCTGGCAGCTAAGATGTATTTGGAACAGTACA 4762  
QY 4667 AGATCTAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGAGCTGGAT 4726  
Db 4763 ACATCTAAGAGTAAAGTTTGGTAAAGGCTACTTTTGGAAATTAATTTGAAGAGCTGGAT 4822  
QY 4727 AGAAAACCTAGAAAGTAGACCGCCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCAAG 4786  
Db 4823 AGAAAACCTGGAATAGATCGCCTTCAAAGAGAAATTCATATATTTTCCCAATGCAAG 4882  
QY 4787 CCGTCAAGAAAGTTTTTCTTATTTTGGCTTATAAAATTCCTAAGAGAGATGTTCAAGTC 4846  
Db 4883 CCGCCCAAGAGAGCTTTTCTTATTTTGGCTTTTAAATTTCCAAAAGAGAGATGTTCAAGTC 4942  
QY 4847 CCTTTTCAATCTTTTTTAAAGCTGAAGAGCTAAACATGCTTTTGGCCATTTGAAGAATA 4906  
Db 4943 TCTCTCAATCTTCTGCTAAGTTGGAAGAGCAAAACGCATTTTGGCCATTTGAAGAATA 5002  
QY 4907 TAGCTTTTCTCAAGCAACATGGAACAGGTTTTTGTAGAACTCTCAATAAGAACAGAGGA 4966  
Db 5003 CAGCTTTTCTCAAGCAACCTCGAAACAGGTTTTTGTAGAACTTACTTAAAGAGCAAGAGGA 5062



QY 4967 GGAAGATAATAGTTGGAACTTTAAACAGCACACATTTGGTGGGAACGAACACACAGAAGA 5026  
 Db 5063 AGAAGATAACAGCTGGGACCTTTGGCCAGCACTCTCTGGTGGGAAGAGGACGAGGAGA 5122  
 QY 5027 TAGAGTAGATTTTGAATTTGAT-----TGTTGGTGTGCT 5063  
 Db 5123 TAGAGTAGATTTTGAATTTGAT-----TGTTGGTGTGCT 5182  
 QY 5064 TACTGGGACTCTCTCTTTTTCACCTTAATTTAACTTTGGTTTAAAGTTTTTATTGG 5123  
 Db 5183 TTTTATTATCTATTTTCTCTTTTATCTATCTATCTTTGATTTTAAAGCTTATATTG 5242  
 QY 5124 AATGTAACCTGAGAACCAAGAACCACTTGAATTTTTCCTTAAGCTCTTAAATGAATG 5183  
 Db 5243 AATAGTTA-TTAAAAAATAAGAACCACTTGAATTTTCTTAAAGCTTAAATTAACACA 5301  
 QY 5184 CTGTGGTGTGTTGTTGCTTTCTTTTAAATAAAGCACTATGTA 5226  
 Db 5302 CCATGGTTATGTTGGATTTCTTTCAATAATTTGCGTGA 5344

RESULT 12  
 RNO426052  
 LOCUS RNO426052 5170 bp mRNA linear ROD 08-JAN-2003  
 DEFINITION Rattus norvegicus mRNA for Abca5 protein.  
 ACCESSION AJ426052  
 VERSION AJ426052.2 GI:27368658  
 KEYWORDS abca5 gene; ATP-binding cassette protein.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE  
 AUTHORS Petry, F., Kotthaus, A. and Hirsch-Ernst, K.I.  
 TITLE Cloning of human and rat ABCA5/Abca5 and detection of a human  
 splice variant  
 JOURNAL Biochem. Biophys. Res. Commun. 300 (2), 343-350 (2003)  
 MEDLINE 22392638  
 PUBMED 12504089  
 REFERENCE  
 AUTHORS Petry, F.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-DEC-2001) Petry F., Department of Toxicology,  
 Georg-August-University Goettingen, Robert-Koch-Strasse 40, Lower  
 Saxony, D-37075 Goettingen, GERMANY  
 revised by author [27-AUG-2002]  
 FEATURES  
 Location/Qualifiers  
 1..5170  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="Wistar"  
 /db\_xref="taxon:10116"  
 1..5170  
 /gene="abca5"  
 36..4964  
 /gene="abca5"  
 /function="transporter"  
 /codon\_start=1  
 /evidence=experimental  
 /product="ATP-binding cassette protein 5"  
 /protein\_id="CAD19800.2"  
 /db\_xref="GI:27368659"  
 /db\_xref="GOA:Q8CF82"  
 /db\_xref="SPTREMBL:Q8CF82"  
 /translation="MATAIRDVGVWQTRTLLKNVLKCRKKSSVQELPPLFPLF  
 WLLISWMHPNKYEVSVDIELSPMDKSLNSLILGYTPVNTTSSVMQVSDHLPD  
 VLVTEAYASKEELASLSKPSNFVGVKDVMSLEKFPFDPVPSVVSFVMSRAGCS  
 KSCDAQWSEGFALQASDAALIQKTNVSLWLESTKAVIMGEAAVEIDTPFR  
 GVILYVIAFSPFGYLAHVAERKLEKLEKLMGLDHTAFWLSWLLYLSLIFL  
 MSLLMAYATASSLPQSSSIVIFLEFLYGLSVPFALMLTFLFKSKHVGVEPFF  
 TWFGVGLLVLVESPRSLVWFLPCCOAFILGIAQWMLDFNEGALFSLREG  
 PPLTITLMLALDSVYALLAVLDQVIFGFRSSLYFLKPSWMSKNKRYKEL

gene

CDS

## ORIGIN

Query Match 70.9%; Score 3879.8; DB 10; Length 5170;  
 Best Local Similarity 85.3%; Pred. No. 0;  
 Matches 4385; Conservative 0; Mismatches 732; Indels 26; Gaps 4;  
 QY 97 AGGTTTATTCAGAAAAATGTCCTCAATTTAGGAGGTAGGAGTTTGGAGACAGACCA 156  
 Db 19 AAGCTGACCTAGAAAAATGCTACTGCAATTTAGGATGTGGCGTTTGGAGACAGACCA 78  
 QY 157 GAACCTCTTCTAGAGATTACTTAATAATGAGACCAAAAGAGTAGTGTTCAGG 216  
 Db 79 GAACCTCTTCTAGAGAAATACCTCGTCAATGAGGAGCTAAAGAGTAGTGTTCAGG 138  
 QY 217 AAATCTTTTCCACTATTTTTTTTATTTGTTGTTAATTAATTTAGCATGATGCATCAA 276  
 Db 139 AAATCTTTTCCCTCTATTTTCTTATTTGCTGATATTAATTTAGCATGATGCATCAA 198  
 QY 277 ATAAGAAATATGAAGAGTGCCTAATATAGAACTCAATCTTATGACAAAGTTTACTCTTT 336  
 Db 199 ATAAGAAATATGAAGAGTATCTGATATAGAACTCAGCCCTATGACAAATCCATCTTT 258  
 QY 337 CTAATCTAATCTTGGATATCTCCAGTGACTAATTTACAAGCAGCATCATGCAGAAAG 396  
 Db 259 CCAACCTTATCTGGGTACACTCCAGTGACTAATTAACAAGCAGCGTATGCAGAGAG 318  
 QY 397 TGTCTACTGATCATCTACCTGATGCTAATTTACTGAAGAATATACAAATGAAAAAGAAA 456  
 Db 319 TTTCTACCGATCATCTTCCGACGTTCTAGTTACTGAAGAAATACGCCAGTGAGAGGAGC 378  
 QY 457 TGTAAACATCCAGTCTCTTAAGCCGAGCAACTTTGTAGTGTGTTTCAAGACTCCA 516  
 Db 379 TGTGGCATCCAGTTTGTCTAAGCCAGCAATTTCTGGGTGTGTTTCAAGAGCTCA 438  
 QY 517 TGTCTTATGAACCTCGTTTTTTTCTCTGATATGATTTCCAGTATCTCTTATTTATATGAT 576  
 Db 439 TGTCTTATGAACCTCGTTTTTTTCTCTGATATGATTTCCAGTGTCTCTTATATATGAT 498  
 QY 577 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGCTCAGGTTTTCACAG 636  
 Db 499 CAAGAGCTGGCTGTTCAAAAGTATGTGATGCTGCTCAGTACTGCTCAGGTTTTCACAG 558  
 QY 637 TTTTCAAGCATCCATAGATGCTGCCATTATACAGTTGAAGACCAATGTTTCTCTTGA 696  
 Db 559 CTTTACAGGCTCAATAGATGCTGCCATTATACAGTTGAAGACCAATGTTTCTCTTGA 618  
 QY 697 AGAGCTGAGTCAACTAAAGCTGTTATTTGGGAGAACTGCTGTTGTAGAAATAGATA 756  
 Db 619 GCGAGCTGGAGTCAACCAAGCTGTCAATCATGGGGAGAGCTCGGTCGTGGAGATTGACA 678  
 QY 757 CTTTCCCCGAGGAGTAAATTTTAAATATACCTAGTTTATAGCAATTTTCCACTTTGGATCT 816

Db 679 CCTTCCACGAGGGTCATCCTCATCTACCTCGTCATAGCCCTTCGCGCCCTTCGGGTACT 738  
Qy 817 TTTTGGCAATTCATATCGTAGCAGAAAGAAAGAAATAAAGAAATTTTAAAGATAA 876  
Db 739 TTCTGGCAATTCACATTTAGCAGAAAGAAAGAAAGAGGTAAAGAGATTTTAAAGATAA 798  
Qy 877 TGGGACTTCATGATACTGCTTTTGGCTTTTCCCTGGGTTCTCTATATACAAAGTTTAAATTT 936  
Db 799 TGGGACTTCACGACACTGCTTTTGGCTGTCTCGGTTCTCTGTGTACAGGATTTGATCT 858  
Qy 937 TTCTTATGTCCTTCTTATGGCAGTCATTTGGCAGTCATTTGGCAGTCATTTGTTATTTCTTAAAGTA 996  
Db 859 TCCCTATGTCCTTCTTATGGCGCTCATTTGCAACAGCGTCTCGTTATTCCTCAGAGTA 918  
Qy 997 GCAGCAATTCATGATATTTCTGCTTTTTCCTTTTATGGATTTATCATCTGTATTTTGTCTT 1056  
Db 919 GCAGCAATTCATGATATTTTCTGCTTCTTCTTCTGTCAGGATTTGTCATCTGTGTTTTGCTT 978  
Qy 1057 TTAATGTCGACCTCTTTTAAATAATCAAAACATGTCGGATAGTGTAAATTTTGTGA 1116  
Db 979 TAAATGTTGACGCTCTTTTAAATAATCAAAACATGTCGGATAGTGTAAATTTTGTGA 1038  
Qy 1117 CTGTGGCTTTTGGATTTATTTGGCTTATGATAATCTCTATAGAAAGTTTTCCTCAAAATCGT 1176  
Db 1039 CCGTGGTGTGGCTTTTGGCTTGTGATAGTCTCTGTAGAAAGTTTTCCTCAGGTCG 1098  
Qy 1177 TAGTGTGGCTTTTCACTGCTTCTGTCTACTGTACTTTTGTGATTTGATTTGCAAGGTCA 1236  
Db 1099 TAGTGTGGCTTCTTCACTGCTTCTGTCTACTGTCTGCGCTTTCTGATTTGGATTTGCAAGGTCA 1158  
Qy 1237 TGCATTTAGCAATTTTAAAGAGTGTCTCAATTTTCAATTTTCAATTTGACTCGAGGCCCATATC 1296  
Db 1159 TGCATTTAGCAATTTTCAATGAGAGCTTATTTTCTAGTTGACTGAGAGGCCCTTACC 1218  
Qy 1297 CTTCAATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTC 1356  
Db 1219 CGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTTTC 1278  
Qy 1357 TCTATCTTCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATTTTTC 1416  
Db 1279 TGTAATTTAGCAATTTTCAATGAGAGCTTATTTTCTAGTTGACTGAGAGGCCCTTACC 1338  
Qy 1417 TGAAGCTTTCAATTTTGGTCAAAAGAGCAAAAGAAATTTATGAGAGTTATCAGAGGGCAATG 1476  
Db 1339 TAAAGCCATCATATTTGGTCAAAAGAGCAAAAGAAATTTATGAGAGTTATCAGAGGGCAACA 1398  
Qy 1477 TTAATGGAATTAATTTAGTTAGTGAATTTATGAGCCAGTTTCTTCAGAAATTTGTAGAA 1536  
Db 1399 TCAATGGGAATTAATTTAGTCTCAATGAATTTGTTGAGCCCGTTTCTTTCAGAAATTTATAGGAA 1458  
Qy 1537 AAGAAGCCATAAGAAATTTAGTGGTATTCAGAAGACATACAGAAGAGGGGTGAAATTTGG 1596  
Db 1459 AAGAAGCTATAAGAAATTTAGTGGTATTCAGAAGCCATATAGGAAGAAATGAGACCGTGG 1518  
Qy 1597 AGCTTTGAGAAATTTGCTTATGACATATATGAGGGTCAGATTTACTGCTTACTGCTTACTGCTT 1656  
Db 1519 AGCTTTGAGAAATTTGCTTATGACATATATGAGGGTCAGATTTACTGCTTACTGCTTACTGCTT 1578  
Qy 1657 ACAGTGGAAACAGAAAGAGTACATTTGATGAATTTCTTTGAGACTCTGCGCCACTTCTG 1716  
Db 1579 ACAGTGGAAACAGAAAGAGTACATTTGATGAATTTCTGTTGAGACTCTGCGCCACTTCTG 1638  
Qy 1717 ATGGGTTTGCATTTATATATGAGCAACAGAGTCTCAAGAAATGATGAATTTGATGAGCA 1776  
Db 1639 ATGGGTTTGCATTTATATATGAGCAACAGAGTCTCAAGAAATGATGAATTTGATGAGCA 1698  
Qy 1777 GAAATATGATGCGATTTTCTCAGAGTTAGATATACATTTGATGTTTGTGAGTGTGAGTGTGAG 1836  
Db 1699 GGAATATGATGCGATTTTCTCAGAGTTAGATATACATTTGATGTTTGTGAGTGTGAGTGTGAG 1758  
Qy 1837 AAAATTTTCAATTTTGGCTTCAATCAAGGGATACCAAGCCCAATATATATACAGAAAG 1896  
Db 1759 AAAATTTTCAATTTTGGCTTCAATCAAGGGATACCAAGCCCAATATATATACAGAAAG 1818

Qy 1897 TGCAGAGCTTTTACTAGATTTAGACATCGACAGCTATCAAGATACCAAGCTTAAAAAT 1956  
Db 1819 TACAGAAAGTTTACTTGGATTTGGACATGCAAGCTATCAAGATATCAAGCGAATAAAC 1878  
Qy 1957 TAAAGTGGTGGTCAAAAGAAAGAGTGTCTATTAGGAATTCCTGTTTGGGAACCCAAAGA 2016  
Db 1879 TAAAGTGGTGGTCAAAAGAAAGAGTGTCTATTAGGAATTCAGTTCTTGGGAATCCAAAGA 1938  
Qy 2017 TACTGTGCTAGATGAATCAACAGCTGGGAATGGACCCCTGTTCTCGACATATTTGATGGA 2076  
Db 1939 TCTCTGCTAGATGAGCCGACAGAGGAATGACCCCTGCTCTCGCCACATTTGTTTGA 1998  
Qy 2077 ATCTTTTAAATAACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTCTAGGATG 2136  
Db 1999 ATCTTTCTCAAGTACAGAAAGGCTTAACCGAGTGACAGTCTTTCAGCACTCACTTCATGATG 2058  
Qy 2137 AGCTGACATCTTTCAGATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTGTTG 2196  
Db 2059 AGCCCGACATTTTTCGCGACAGGAAGCTGTCTATATCAAGGAATGCTGAAATGTGTTG 2118  
Qy 2197 GTTCTTTCAATTTCTCAAAAGTAAATGGGGATCGGCTACCCCTGAGCATGTACATAG 2256  
Db 2119 GTTCTTTCAATTTTCTGAAAGTAAATGGGAATCGGCTACCGCTCTGAGCATGTATATAG 2178  
Qy 2257 ACAAATATGTCGCCACAGATCTCTTTCTTCTACTGCTTAAACAATATACCTTGGAGCTA 2316  
Db 2179 ACAGGTACTGTGCCACAGAGTCTGCTCTGCTGCTGTAGACAGACATCCCTGCGGCG 2238  
Qy 2317 CTTTATTAACAAGATGACCAACAACTTGTGTATAGTGTGCTTTTCAAGGACATGAGCA 2376  
Db 2239 CTCTGTCAGCAGAAATGACAGCATTTGTCTATAGCTTGGCTTCAAAAGACATGAGCA 2298  
Qy 2377 AATTTTCAGTTTGTGTTTCTGCGCTTACAGTCATTTCAATTTGGGTGTCATTTCTTATG 2436  
Db 2299 AATTTTCAGCTTGTGTTTCTGCTCTAGACATTTCAAACTCTGGGTGTAATTTCTTATG 2358  
Qy 2437 GTCTTTCCATGACCATTTTGGGAAGCTATTTTAAAGCTAGAAGTTGAAGCAGAAATTTG 2496  
Db 2359 GTCTTTCCATGACCATTTTGGGAAGCTGTTTTTAAAGCTAGAAGTTGAAGCAGAAATTTG 2418  
Qy 2497 ACCAAGCAGATTTATGATGTTTATCTCAGACGCCATCTGAGGAGGAATTTGATTTCAAAAT 2556  
Db 2419 ACCAAGCAGATTTATGATGTTTATTTACCCAGCAGCCAGGAGGAAGAAACAGATTTCAAAGT 2478  
Qy 2557 CTTTTCATGAAATGGAACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTTCTCTAGTGA 2616  
Db 2479 CTTTTCATGAAATGGAACAGAGTTTACTTATTTCTCTGAAACCAAGCTTCTCTAGTGA 2538  
Qy 2617 GCACCATGAGCCCTTTGGAAACCAACAGATGTATCAATAGCAAGTTTCAATTTCTTACCT 2676  
Db 2539 GCACCATGAGCTCTCTGGAAGCAGCAAGTATCTACAAATAGCAAGTTTCTTCTCTCAC 2598  
Qy 2677 TGAACGTGAAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736  
Db 2599 TGAAGCGAAGAAACAAATCAGTGGCTCTGTGTGCTTCTGCTTTTAAATTTTTCAG 2658  
Qy 2737 TTCAGATTTTATGTTTTTGGTTTCATPACTCTTTTAAATAATGCTGGTTCATCAAC 2796  
Db 2659 TTCAGATTTTATGTTTTTGGTCCATCATTTTAAATAATGCTGGTTCATCAAC 2718  
Qy 2797 TTGTTCCAGCTTATTTTCTTAAACCTGGAGCAACCAATCAATCAACCAAGTC 2856  
Db 2719 TCGTTCCAGACTTGTATTTTCTTAAACCTGGAGTAACCTCAATAATCAACCAAGCC 2778  
Qy 2857 TGTCTTCTTCAAAATTCCTGCTGACTCAGATATCAGTGTATTTAGTCTTTTTCACAGCC 2916  
Db 2779 TGTCTGCTTCAAAATTCCTGCTGACTCAGATATCAATGATCTTATTTGACTTTTTCACAGCC 2838  
Qy 2917 AGAACATATGTCGAGCATGATTAATGACAGTCACTATGATCCGTGGCTCCCCATAGT 2976  
Db 2839 AGAACATATTTGGCCATGTTTAAATGACAGTCACTATGCTGCTGCTCTCTCATAGT 2898

QY	2977	CGGCTTTAAATGTGATGCAATTCAGAAAAGGACTATGTTTTTGGCAGCTGTTTTTCAACAGTA	3038
Db	2899	CGGCTCTCAATGTGGTGAGTCAGAAAAGGACTATGTTTTTACTGCTGTTTTTCAACAGTA	2958
QY	3037	CTATGGGTTTATCTTTTACCTATATTTAGTGAATATCATTAGTAACCTACTATCTTTATCATTT	3096
Db	2959	CTATGGGTTTATCTTTTGGCAGTAATGATGAACATCATTAGTAACCTACTATCTCTATCATTT	3018
QY	3097	TAAATGTGACGTGAAACCAATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTTACTGATA	3156
Db	3019	TAAATGTGACGTGACCACTATCCAGATCTGGAGTACCCCGTTTCATTTCAGAAAATTTACTGACA	3078
QY	3157	TAGTTTTTAAAAATTGAGCTGTATTTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTTACTGCAA	3216
Db	3079	TTGTTTTTAAAGTTGAGCTATTTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTTACTGCAA	3138
QY	3217	TGCCACCTTACTTTTGCCATGAAAAATGCAGAAATCATAAAGATCAAAAGCTTATATCTCAAC	3276
Db	3139	TGCCACCTTACTTTTGCCATGAAAAATGCAGAAATCATAAAGATCAAAAGCTTACTCAGC	3198
QY	3277	TTAAACITTTCAGGTCITTTTGCCATCTGCATATTGGATTGGACAAGCTGTTGTTGATATCC	3336
Db	3199	TTAAACITTTCAGGTCITTTTGCCATCCGCTACTCGATTGGACAGGCTGTGGTGGACATTC	3258
QY	3337	CCTTATTTTTTATCATTTCTTATTTTGATGCTAGGAAGCTTATGGECATTTTCATTATCGAT	3396
Db	3259	CCCTGTTTTTTGTTGTTCTGACTTTGATCTGGAGATTTTATTTGCTTTTCATCACGGAC	3318
QY	3397	TATATTTTTTATCTGTAAAGTTCCTGCTGTGGTTTTTTTGGCTTATTTGGTTATGTTCCAT	3456
Db	3319	TGTATTTTTTATCCGTGTAAGTTTCTTGCTGTGGTTTTTTTGGCTCAITTGCTTATGTGCGCT	3378
QY	3457	CAGTTATCTGCTTCACCTTATATGCTTCTTTTTCACCTTTAAGAAAATTTTTAAATACCAAG	3516
Db	3379	CGGTTATCTGTTTCACTTATATAGTTCGTTTTCACCTTTAGAAAATTTTTTAAATACCAAG	3438
QY	3517	AATTTTGGTCAATTTATCTTATCTGTGGCAGCGTTGGCTTGTTATGCAATCACTGAAATAA	3576
Db	3439	AAATTTGGTCAATTTATCTTATCTGTGACAGCATTTGGCTTGTTGTCGCACTCACGGAAATAA	3498
QY	3577	CTTTCTTTTATGGGATACACAAATTCGAATTTCTTCTCATATATGCGCTTTTGATCATCATTC	3636
Db	3499	CTTTCTTTTCTGGGTATGGAGTTAGCGTGTTTTTTTCATTCACCTTCTGCAATGCCATTC	3558
QY	3637	CAATCTATCCACTTCTAGGTTGCGCTGATTTCTTTTCATAAAGATTTCTTGAAGAATGTAC	3696
Db	3559	CAATCTACCGCTCCTGGTTGTCTGATTTCTTTTCATAAAGGTTCTTGAAGAATATATAC	3618
QY	3697	GAAAAATGTGGAACCACTATTAATCATGGGATAGCTTTTCAGTAGCTGTTTATATCGGCTT	3756
Db	3619	CAAAAACTGAGAATGCCATACATCCCTGGGACAGACTTTTATGTTGCTGTATCATGCGCTT	3678
QY	3757	ACCTGCAGTGTACTGTGGATTTTCCCTTTTACAATCTATGAGAAAAAATATGGAGGCA	3816
Db	3679	ACCTGCAGTGTACTGTGGATTTTCCCTTTTACAACATATGAGAAAAAATATGGAGGCA	3738
QY	3817	GATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTAAAAATAGGAAGC	3876
Db	3739	GATCAATAAGAAAAAGATCCGCTTTTTCAGGGCCCTCTCACAAAAAGCCAAACATTAAGAAT	3798
QY	3877	TTCCAGAACCAACAGACAATCAGAGATGAAGATGAAGATGTCAAAGCTGAAAGACTAAAGG	3936
Db	3799	TTCCAGAACCAACCACTAATGAGGATGAAGATGAAGATGTCAAAGCGGAAGGTTGAAGG	3858
QY	3937	TCAAAGAGCTGATGGGTTGCCAGTGTGTGTGAGGAGAAACCATCCATTTATGGTCAGCAATT	3996
Db	3859	TTAAAGAGCTGATGGGTTGCCAGTGTGTGTGAGGAGAAACCGGCCCATTTATGCTATACAAAT	3918
QY	3997	TGCATAAAGATATGATGACAGAAAGATTTTCTTTCTTCAAGAAAAAGTAAAGAAAGTGG	4056
Db	3919	TGCATAAAGATATGATGACAGAAAGATTTTCTTTCTACTCAAGGAAAAACACAAAGATAG	3978
QY	4057	CAACTAAATACATCTCTTTCTGTGTGAAAAAAGGAGAGATCTTAGGACTATTTGGGTCCAA	4116

3979	Db		CAACAAATACGTCCTCTTCTGTGTGAAAGGAGAGATCTTTGGGACTGTTGGGTCCAA	4038
4117	Qy	ATGGTGTGCGAAAGCACAAATTAATAATCTGTGGTGGTGATATTGAACCAACTTCAG	4176	
4039	Db	ATGGTGTGCGAAAGCACGATTATTAATCTTCTGTGGGATGTTGAACCAACTTCAG	4098	
4177	Qy	GCCAGGTATTTTATAGAGATTAATCTTTCAGAGACAAGTGAAGATGATGATCTACTGAACT	4236	
4099	Db	GCAAGATATTTCTAGGAGATTAATGGTTTCATTTCAAATGAAGATGATGAATCACCAGAT	4158	
4237	Qy	GTATGGTTACTCTCTCAGATATAACCCCTTTGTGGCCAGATCTACTCATTTGCAGGAACAT	4296	
4159	Db	GTATGGTTACTGCCCCAGACAAACCCACTCTGGCCAGATATTACTCTGCAAGAACAT	4218	
4297	Qy	TTGAAATTTATAGAGCTGTCAAAAGGAATGAGTGCAAGTGACATGAAGAAGTCTAATGTC	4356	
4219	Db	TTGAAATTTATGAGCTGTGAAAGGAGATGAGTTCAAGTGACATGAAGGAAGTAAATCAGTC	4278	
4357	Qy	GAATAACACATGCATTTGATTTAAAGAACATCTTCAGAAGACTGTATAAGAAACTACCTG	4416	
4279	Db	GAATAACAAAGCCCTTGATTTGAAGAACATCTTCAGAAACTGTAAAGAACTCCCTG	4338	
4417	Qy	CAGGAATCAACGAAAGTTGTGTTTTGCTCTAAGTATGCTAGGAATCTTCAGATTAATT	4476	
4339	Db	CAGGGATCAAGCGAAAGCTGTGTTTTGCTCTCAGTATGCTGGGAATCTCTCAGGTGACAC	4398	
4477	Qy	TGCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAACAGCACATGTCGGGAGCAA	4536	
4399	Db	TGCTGATGAGCCGCTCTACAGGTATGGACCCCAAGCCAAACAGCACATGTCGGGAGCCA	4458	
4537	Qy	TTCGAACTGCATTTAAAAACAGAAAGCGGCGCTCTATTCTGACCCTCACTATATGAGG	4596	
4459	Db	TTCGAACTGGTTTTAAAAACAAGACGGGCTGCCCTCTCTCACTACGCATTCATGGAAG	4518	
4597	Qy	AGCCAGAGCGCTGCTGTGATCGAGTACGTATCATGTGTCTGGCGCAGTTAAGATGATCG	4656	
4519	Db	AGCCGAGGCGTGTGTGACAGGGTGCCCATCATGTGTCCGGCGAGCTAAGATGCAATTG	4578	
4657	Qy	GAACAGTACAAACATCTAAAGAGTAAATTTGGAAAAGGCTACTTTTTGGAAATTAATTTGA	4716	
4579	Db	GAACAGTACAAACATCTAAAGAGCAAGTTTGGCAAGGCTACTTTTTAGAAATTAACATAA	4638	
4717	Qy	AGGACTGGATAGAAAACCTTAGAGTAGACCGCCCTTCAAAGAGAAATTCAGTATATTTTCC	4776	
4639	Db	AGGACTGGATAGAAAACCTTGGAAATAGATGCGCCTTCAAAGAGAAATTTCAATATATTTTC	4698	
4777	Qy	CAAAATCAAGCCGCTCAGGAAAGTTTTTCTTCTATTTTGGCTTATAAATTCCTTAAGGAAG	4836	
4699	Db	CAAAATCAAGCCGCGCAAGAGAGCTTCTTCTCTATTCTGGCTTATAAATTCCAAAGAGAG	4758	
4837	Qy	ATGTTCACTCCCTTTCAAACTTTTTTTAAGCTGGAAGAGCTAAACATATCTTTTGGCCCA	4896	
4759	Db	ATGTTCACTCTCTCTCACAGTCCCTTGCTAAGTTGGAGGAAGCAAAACACATTTTGGCCA	4818	
4897	Qy	TTGAGGAATATAGCTTTTCTCAAGACACATGGAACAGGTTTTTGTAGAACTCACTAAAG	4956	
4819	Db	TTGAGGAATACAGCTTTTCTCAAGACACCTCGGAACAGGTTTTTGTAGAACTTACTAAGG	4878	
4957	Qy	AACAAGGAGGAGGAATAATAGTTGTGGAACTTTAAACAGCACACTTTGGTGGGAACGAA	5016	
4879	Db	AGCAGGAGGAGGAAGATAATAGCTGCGGACTTTTAAACAGCACGCTCTGGTGGGAAGAA	4938	
5017	Qy	CACAGAGATAGAGTAGTATTTTGAATTTGT--ATTGTTGGGTCTGCTTACTGGGACT	5073	
4939	Db	GGCAGGAAGACAGAGTCGTGCTCTGAATGCTCGGCTCACCCGCTGCTGACCCCTGAAC	4998	
5074	Qy	TCCTTTCTTTTCA-----CTTAATTTTAACTTTGGTTTAAAAAGTTTTTTT	5118	
4999	Db	TTACTCTTTTATTATTGTTTTTCTCTCTTTTATCTTAACTTTGTTTTTAAAAAGTTTATT	5058	
5119	Qy	ATTGGAATGTTAACTGGAGAACCAAGAACGCACTTGAAATTTTCTAAGCTCCTTAATTC	5178	

Db 5059 ACTGATAGTACAC-----AAACATGAAGACCTTGGAATTCCTAACTCCTTAATTT 5114

QY 5179 AAATGCTGTGG-----TTGTGCTTTTGTCTTTCTTTTAAATAAA 5217

Db 5115 ATGCACCATGGTTATTATGCTGCTGGGTTTCTTTTAAATAAA 5157

RESULT 13

AB097675

LOCUS

DEFINITION Mus musculus abcas mRNA for ABC transporter subfamily A ABCA5, complete cds.

ACCESSION AB097675

VERSION AB097675.1 GI:29420876

KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS Kubo, Y., Satsuma, Y., Sekiya, S., Nada, S. and Yamaguchi, A.

TITLE Molecular cloning of mABCA5, the mouse homologue of ABCA5

JOURNAL Unpublished

REFERENCE Kubo, Y., Satsuma, Y., Sekiya, S., Nada, S. and Yamaguchi, A.

AUTHORS Direct Submission

TITLE Submitted (06-DEC-2002) Yoshiyuki Kubo, I.S.I.R., Osaka Univ.;

JOURNAL Mihogaoka 8-1, Ibaraki City, Osaka 567-0047, Japan

(E-mail: kubos3@sanken.osaka-u.ac.jp, Tel: 81-06-6879-8546, Fax: 81-06-6879-8549)

FEATURES

source

1..4929

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/chromosome="11"

/map="11E1"

/tissue\_type="brain"

/dev\_stage="newborn"

1..4929

/gene="abcas"

1..4929

/gene="abcas"

/codon\_start=1

/product="ABC transporter subfamily A ABCA5"

/protein\_id="BAC6658.1"

/db\_xref="GI:2942087"

/translation="NATAIDYGVWQRTRILLKNYLIKRTKSSVQEIPLPFLFLP  
WLVSMHPNKKYEVSDIELSPMDKFSLSNVLIGVTPVNTITSSIMQVSTDLHPK  
VVTREYANERKELVAASLSKSNFVGVKDTMSVELRFPPEMIPVSIYMSREGGS  
KTCDAQWWSLGTVLQASIDAAIIQLATNVSVMSELESTKAVTMGEAAVVEIDTPR  
GVLLYIATSPFGYFAIHVAEKEKLEKFLKINGLHDTAFLSNLLYASLIFL  
MSLMVATATASSLPQSSSIVIFLLFYLGLSSVFPALMLTLPFKSKHGVGVVEFFV  
TVVFGVGLLAVLIESFRLSVLWFLSPQCQAFGLTGAQVHLEDNEGALFSLNTEG  
PYLITITIMALSDVFLVLLAVLDVIGFGLRSLSYFLPKPSYKKNRYKSL  
SGNTINGNISLNEIVPSSBFIKEAIRISGIVSKYKKTENVEALRNLSFDIYEQQ  
ITALLGHSCTGSLTNLTCGLCPSSDGFASITGYHRVSEIDEIPEAKKIGICPSDI  
NFDVLTPENLSILASIKGIPANNIIQEVQVLLDLDMQAKDKQAKLSGQKRKLS  
VGAVLTPKILLDEPTAGMDPCSRHIVNLLKRYKARVTVFSTHPEADTLAOR  
KAVISGMLKVGMSDFILKSKWIGIRLSMVIDRYCATESLSSIVRQHIIPAAULQON  
DQQLYSLPFLKMDPSGLFSDALDHSNLGIVSYGSMNTLLEVLKLEBAEIDQAD  
YSVFTQCPREEDTSKSDMEQSLILSETKASVSTMSLWQQVSTIAKPHLSLK  
RESKSVRAVLLALLIFPAVOIFMFLHHSFKNAVPIKLVPLDYFLKPGDKPHKYLS  
LLQNSDTNGLITFPAHONIMVAFNDSDYVSAAPHSAAANVRSKDYVPSAVP  
NSTMYCIPWNNII SNYXLHLYNVAETAIQTWSTPFIQIETIDIVFKLELYFOAALIG  
IVTAMPPFAMENAEHKKI KAYQLKSLGLPSAYWQAVVDIPLPVLILMLGSL  
FAFHGLYIFPAKFAVFLVAVFVPSVILFYIATFTFKILNTKFWFSIYVTL  
ACVATETTFYQYATVAVHVTFCIAPIYPLGLCLISFIKSGMKNPNENTINPW  
DRLLVAVNYQCLIMWIFLLQHYEKIHGSRINRDPFRALSQAKNKKPEPPINE  
DEEDVKAERLVCXELMGCCQCEKPAIMVGNLHKEYDDKDFLHRSKRTKVAKYIS  
FCVKEIILGLLPGNAGKSTVINTLVGDEPTSGIKIPLGIDYGSSEDDESICKMGY  
CQQTWPLPDLTFLGHFEIYGAVKGSPQDKVENISRTIKALDLKHLQKTVKLPAG  
IKKLCFALSLMGNQVILLDBPSGMDPRAKQHWRAIRTAFKNKKRAALLTTHYME

ORIGIN

Query Match 70.7%; Score 3873; DB 10; Length 4929;

Best Local Similarity 86.6%; Pred. No. 0;

Mismatches 0; Matches 660; Indels 0; Gaps 0;

114 ATGTCACATGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGAACACTTCTTCTAGTAAG 173

Db 1 ATGGCTACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGAACACTTCTTACTGAAA 60

QY 174 AATTACTTAATTAATGACAGAACCAACCAAGAGAGTAGTGTTCAGGAAATTCCTTTTCCACTA 233

Db 61 AATTACTTAATTAATGACAGAACCAACCAAGAGTAGTGTTCAGGAAATTCCTTTTCTCTA 120

QY 234 TTTTCTTATTTTGGTTAATTAATTAATGACATGATGATCCAAATAAGAAATATGAAGAA 293

Db 121 TTTTCTTATTTTGGCTGATATTAGTTAGCATGATGATCCAAATAAGAAATATGAAGAG 180

QY 294 GTGCCTAATTAAGACTCAATCCTATGGACAAAGTTTACTTCTTCTAATCTAATCTTCTGGA 353

Db 181 GTATCTGATATAGCTCAGCCCTATGGACAAATTCAGCCTTTCACACGTTATCTTCTGGA 240

QY 354 TATATCTCCAGTGAATAATTAACAAGCAGCATCATGACAGAAAGTGTCTACTGATCATCTA 413

Db 241 TACACTCCCGTGACTAACTTACAAGCAGCATTATGACAGAGGTTTCTTACCGATCATCTT 300

QY 414 CTTGATGTCATTAATTAATGACAGAAATATAAATAAGAAAGAAATGTTAATCATCCAGTCTC 473

Db 301 CCAAGCTTATAGTTTACTGAAGAAATACGCAATGAGAAAGAACTGGTAGCCGCAAGTCTT 360

QY 474 TCTAAGCCGACCAACTTTCTAGTGTTGTTTCAAAGACTCCATGCTCTATGAACCTTCT 533

Db 361 TCTAAGTCAGCAACTTCTGAGTGTTGTTTCAAAGACTCCATGCTCTATGAACCTTCT 420

QY 534 TTTTCTTCTGATATGATCCAGTATCTTCTTATTTATGATGATCAAGACTCGCTGTGTTCA 593

Db 421 TTTTCTTCTGAAATGATCCAGTGTCTTCTTATTTATGATGATCAAGAGAGGCTGTCTCA 480

QY 594 AAATCATGTGAGGCTCTCAGTACTGCTCTCAGGTTTTCAGGTTTTCACAGTTTTCACAGCATCCATA 653

Db 481 AAGACATGTGATGCTGCTCTCAGTACTGCTCTTTGGGGTTTACAGTTCTGCGAGCATCGATA 540

QY 654 GATGCTGCCATATATACAGTTGAGCAACTGTTTCTTCTTGGAGAGAGCTGGAGTCACT 713

Db 541 GATGCTGCCATATATACAGTGAAGCAACTGTTTCTGTGTGAGGAGCTGGAGTGCACC 600

QY 714 AAAGCTGTTTATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCCGAGGAGTA 773

Db 601 AAAGCTGTGATCATGGAGAGGCGCTGTTGTGAGATTGACACCTTCCCGGAGGGGTC 660

QY 774 ATTTTAAATATACCTAGTTATAGCATTTTCACTTTTGGATCTTTTGGCAATCATATC 833

Db 661 ATCTCATCTACCTCTGCTAGCTTCTCGCCCTTCGGCTACTTCTCGCAATTCACATC 720

QY 834 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGACTTCATGATATCT 893

Db 721 GTGCGAAGAAAAAGAAAGAAATTTTAAAGAAATTTTAAAGATAATGGACTTCATGACAT 780

QY 894 GCCTTTTGGCTTTTCCGGGTTCTCTATATACAAAGTTTAAATTTTCTTATGTCCTTCTT 953

Db 781 GCCTTTTGGCTTTTCCGGGTTCTCTGTCAGCAAGCTTGAATTTCTTATGTCCTTCTT 840

QY 954 ATGCGAGTCAATTCGAGCAGCTTCTTTGTTTATTTTCTCAAAGTAGCAGCATTTGTGATATTT 1013

Db 841 ATGCGTGTATCCCAACAGCTTCTTCTGTTATTTCCCTCAGAGTAGCAGCATTTGTGATCTTT 900

QY 1014 CTGCTTTTCTTTTCTTTATGATATCATCTGATTTTCTTTTCTTTTCTTAAAGTAGCAGCATCTT 1073

Db 901 CTACTGTTCTTCTTTATATGATTTGATCTGTTGTTTCTTTGCTTTTAAATGCTGACGCTCTT 960





FEATURES		Location/Qualifiers
source	1..4945	GERMANY
gene	1..4945	/organism="Rattus norvegicus"
CDS	9..2786	/mol_type="mRNA"
		/strain="Wistar"
		/db_xref="taxon:10116"
		/tissue_type="testis"
	1..4945	
	/gene="Abca5"	
	/gene="Abca5"	
	/note="splice variant V+16"	
	/codon_start=1	
	/evidence=experimental	
	/product="ATP-binding cassette protein 5"	
	/protein_id="CAD80052.1"	
	/db_xref="GI:29170396"	
	/translation="NATAIRDVGVWQRTRILLKNYLKCRKSSVQEIILFPLFFPL	
	WLILSMHPKXKVEVDIIEISPMDSIILNLIIGVTPVNTTSSVMQVSTDLPLD	
	VLVTEYASEKELASSISKPSNFVGVKDVMSYELRFRPDMVPVSVMDSRAGS	
	KSCDAQYWSGFTALQASIDAAIIQLKTNVSLWRELESTKAVIMGEAAVEIDTPR	
	GVLLIYATSPFGFIAIHIVAEKREKRLKEFLKIMGLDTPFLWLSKVVYTLIFL	
	MSLMAVATASSLPQSSIVIFLLFLYGLSSVFFALMTPLPKSKVGVVVEFV	
	TVVFGVLLVLVESFPRSLVWLFSPQCQFLIGIAQVWHLDFNEGALFSLITEG	
	PYPLITLMLALDSVFYALLAVLDQVIGFGLRSLSLYFLKPSYKXKNRYKEL	
	SEGNLINGSINEIYVPSSEPIGEAIRISIOKAVRKNETVEALRNLSFDIYEQG	
	ITALLHSGTSGKSTLNLICGLCPPSDGFASIGHRVSEIDEMFEARKMLICPQSDM	
	NFDVLTVENISILASVIGIPANNIIQEVQVILDLDMQAIKNQAKLSGGQKRLS	
	LGIAVGNPKILLDEPTAGMDPCSRHIVNLLKPKRANRVTFSTHFMDEADILADK	
	KAVISQMKLVGSSIFLKSQWIGYRLSMYIDRYCATESLSLVROHIIPAALLOQN	
	DOQIVYLPFKMDKFSGLFDIHSNLVSYGVSMSTLLEDFYFLKLEVAEIDQAD	
	YSVFTQCPRETSKSDMEQSLILSETKASIVSTMLWKQVSTIAKFHPLSLK	
	RESKSVRSVLLILLITFFAVQIFMPLVHHSFNNAVPIKLVPDLYFLKPGKPKHYKTS	
	LLLQNSGTGSV"	
ORIGIN		
Query Match 69.8%; Score 3820.4; DB 10; Length 4945;		
Best Local Similarity 86.0%; Pred. No. 0;		
Matches 4252; Conservative 0; Mismatches 676; Indels 16; Gaps 1;		
QY	107	AGAAACATGTCACCTGCAATAGGAGGTAGGAGTTTGGAGACAGACAGACACACTCT 166
DB	2	AGAAACATGGCTACTGCAATTAGGAGTGTGGCGCTTTGGAGACAGACAGACACTCT 61
QY	167	ACTGAAGAATTACTTAATAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTCTTT 226
DB	62	TCTGAAGAATTACCTCGTCAATGAGGACTAAAAGAGTAGTGTTCAGGAAATCTCTTT 121
QY	227	TCCACTATTTTATTTTGGTAAATTAATTAATAGCATGATGCATCCAAATAGAATA 286
DB	122	TCTCTATTTTCTTATTTGGCTGATATTAATTAATAGCATGATGCATCCAAATAGAATA 181
QY	287	TGAAGAAGTGCCATAATAGAACTCAATCCCTATGGACAAAGTTTACTCTTTCTAATCTAAT 346
DB	182	TGAAGAAGTATCTGATATAGAACTCAGCCCTATGGACAAATCCATCTTTCCACCTTAT 241
QY	347	TCTTGATATCTCCAGTGAATAATTAATTAAGAGCATGATGCAGAAAGTGTCTACTGA 406
DB	242	CCTGGGTACACTCCAGTGACTAATACTACAGCAGCGCTCATGCAGAGAGTTTCTACCGA 301
QY	407	TCATCTCTACGTATGTCAATAATTAATTAAGAGATATACAAATGAAAAGAAATGTTAAATC 466
DB	302	TCATCTTCCGACGTTCTAGTACTTAAGAAATACGCCAGTGAAGAGTGTCTGGCATC 361
QY	467	CAGTCTCTTAAGCCGACAACTTTGTAGGTGTGGTGTTCCTAAGACTCCATGTCTCTAGA 526
DB	362	CAGTTGTCTAAGCCGACAACTTCGTGGTGTGGTGTTCCTAAGACTCCATGTCTCTAGA 421
QY	527	ACTTCGTTTTTCTCTGATGATGATCCAGTATCTTCTATTTATATGATTCAGAGCTGG 586
DB	422	ACTCGGTTTTTCTCTGATGATGATTCAGTGTCTTCTGTTATATGATTCAGAGCAGG 481
QY	587	CTGTTCAAATAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTTCACAGTTTACAAGC 646
DB	482	CTGTTCAAAGTCTATGTGATGCTGCTCAGTACTGCTCCTCAGGTTTTCACAGTTTACAAGC 541
QY	647	ATCCATAGATGCTGCCATTATACAGTTGAAGACCAATGTTCTCTTTGGAAGGAGCTGGA 706
DB	542	CTCAATAGATGCTGCCATTATACAGTTGAAGACCAATGTTCTCTCTGGAAGGAGCTGGA 601
QY	707	GTCAACTAAAGCTGTTATTATATGGAAGAAACTGCTGTGTGTAGAAATAGATACCTTTCCCG 766
DB	602	GTCAACTAAAGCTGTTATTATGGAAGAAAGTGGCGTCTGGAGATTGACACCTTCCCAAG 661
QY	767	AGGAGTAATTTAATATACCTAGTATTACATTTTCACTTTTGGATATCTTTTGGCAAT 826
DB	662	AGGAGTATCTCTCATCTCTCTCTAGCTTCTGCCCTTCGGGTACTTCTCTGCAAT 721
QY	827	TCATATCGTAGCAGAAAAAATAAAGAAATTTTAAAGATATATGAGACTTCA 886
DB	722	TCATATCGTAGCAGAAAAAATAAAGAGTTTAAAGAGTTTAAAGATATATGAGACTTCA 781
QY	887	TGATATCTGCTTTTGGCTTTCCTGGGTCTCTCTATATACAGTTTAAATTTCTTATGTC 946
DB	782	CGACATCTGCTTTTGGCTGCTCTGGGTCTCTCTGTACAGAGTTGATCTTCTCTATGTC 841
QY	947	CTTCTTATGGCAGTCAATTCGACAGCTTCTTTTGTATTTCCTCAAAGTAGCAGATTGT 1006
DB	842	CTTCTTATGGCAGTCAATTCGACAGCTCTCTGTTATTCCTCAGAGTAGCAGCATGT 901
QY	1007	GATATTTCTGCTTTTCTCTTATGGATATCATCTGTATTTTCTTAAATGCTGAC 1066
DB	902	GATTTTCTGCTCTCTCTCTCTGACAGTTGTCATCTGTGTTTCTTAAATGTTGAC 961
QY	1067	ACCTCTTTTAAAAATCAAAACATCTGGGAATAGTTGAATTTTGTACTGTGCTTT 1126
DB	962	GCCTCTTTTAAAAATCAAGCAGCTGGAGTCTGTGAGTTTGTTCACCGTGTGT 1021
QY	1127	TGATTTATGGCTTATGATTAATCTCATAGAAAGTTTTCCTTAAATGTTAGTGTGCT 1186
DB	1022	TGCTTTTGTGGCTTGTGATAGTCTCTGTAAGAGTTTCCCGAGTCTGAGTGTGCT 1081
QY	1187	TTTCAGTCTCTCTGTCACCTGACTTTTCTGATTTGATTTGACAGAGTCAATGCAATTAGA 1246
DB	1082	CTTCAGTCTCTCTGTCAGTGGCTTTCTGATTTGGATTTGACAGAGTCAATGCAATTAGA 1141
QY	1247	AGATTTTAAAGAGTGTCTTCAATTTTCAAATTTGATGCGAGCCCATATCTCTAATAT 1306
DB	1142	AGATTTTCAATTAAGAGAGCTTATTTTCTAGTTGATGAAAGCCCTACCCGCTAATAT 1201
QY	1307	TACAATATCATGCTCACACTTAATAGTATATTTCTATGCTCTCTTGGCTGTCTATCTGA 1366
DB	1202	CACCTTACCATGCTCGCTCTGACAGTGTGTTCTATGCCCTGCTGGCTGTGTACCTTGA 1261
QY	1367	TCAAGTCAATTCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAAGCTTC 1426
DB	1262	TCAGGTCAATCCCGAGGAAATTTGGCTTGAAGAGTCACTCTTTGTACTTCTAAAGCCATC 1321
QY	1427	ATTTGGTCAAGAGAGAAAAAATATGAGAGTTATCAGAGGCAATGTTTAAATGAAA 1486
DB	1322	ATATTGGTCAAAAAAATAAAGAACTATAAGAGCTATCAGAGGGGCAATCAATGAGAA 1381
QY	1487	TATTAGTTTATGAAATTTATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAAGAGCCAT 1546
DB	1382	TATTAGTCTCAATGAAATTTGTAGCCCGTTTCTTCAGAAATTTATAGGAAAAAGAGCTAT 1441
QY	1547	AGAAATTAGTGTATTTCAGAAACATACAGAAAGAGGGTGAATATGTTGGAGGCTTTGAG 1606
DB	1442	AAGATAAGTGTGTTTTCAGAAAACCTTATAGGAAAGAAATGAGACCGTGGAGGCTTTGAG 1501
QY	1607	AAATTTGTCAATTTGACATATATGAGGTCAGATTACTGCTTACTTGGCCACAGTGGAAAC 1666
DB	1502	AAATTTGTGTTGACATCTATGAAGCCAGATTACTGCTCTGCTGGGCCACAGTGGAAAC 1561
QY	1667	AGAAAGAGTACATTGATGAATATTCTTTGTGGACTCTGCCCCACCTTCTGATGGGGTTGCG 1726



1562 AGGAAAGGACACATTAATGAATATCTCTGTGGAGCTGTGCCACCTTCTGATGGTTGCG 1621  
 1727 ATCTATATATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGAT 1786  
 1622 TTCTATATATGGACACAGAGTCTCTGAAATAGATGAATGTTTGAAGCGAGAAATGAT 1681  
 1787 TGGCAATTTGTCACAGTTAGATATACATTTTGTGTTTGTGACAGTGTAGAGAAATTTATC 1846  
 1682 TGGCATAATGTCACAGTCAGATATGAATCTTTGTGTTCTGACAGTGTAGAGAAATTTATC 1741  
 1847 AATTTTGGCTCAATCAAGGAGTACACAGCAACAAATATAATACAGAGTGTGAGAGGT 1906  
 1742 AATTTTGGCTTCAGTCAAGGAGTACACAGCAACAAATATAATTCAGGAGTGTGAGAGGT 1801  
 1907 TTTACTAGATTTAGACATGTGACATATCAAGAGATAACCAAGCTAAAGAAATTTAAGTGTGG 1966  
 1802 TTTACTGATTTGGACATGTGACATATCAAGAGATAATCAAGGAAATTAATTAAGTGTGG 1861  
 1967 TCAAAAAAGAGTGTCTATTAGGAATTTGCTTTTGGGAACCCAAAGATATCTGTGCT 2026  
 1862 CCAAGAGAGAAAGTGTCTTTAGGAATTCAGATTTCTGGGAATCCAAAGATCTCTCTGCT 1921  
 2027 AGATGAACCAACAGCTGGAATGGACCCCTGTCTCGCATATTTGATGGAATCTTTTAA 2086  
 1922 AGATGAGCCGACAGAGGAATGGACCCCTGTCTCGCCACATTTGTTTGGAACTTCTCAA 1981  
 2087 ATACAGAAAGCAATCGGTTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACAT 2146  
 1982 GTACAGGAAGGCTAACCGAGTGACAGTCTTCAGCACTCACCTTCATGGATGAGGCCGACAT 2041  
 2147 TCTTCAGATAGGAAGTGTGATATCACAGGAATGCTGAAATGTGTGTTGTTCTTCAAT 2206  
 2042 TCTTCCGACAGGAAGTGTGATATCACAGGAATGCTGAAATGTGTGTTGTTCTTCAAT 2101  
 2207 GTTCTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATG 2266  
 2102 TTCTCTGAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTATAGACAGTACTG 2161  
 2267 TGCCACAGAAATCTCTTTCTTCACTGTGTTTAAACAACATATACCTGAGACTATTTATACA 2326  
 2162 TGCCACAGAGTGTCTCTGCTGGCTGGTTAGACAGCACATCCCTCGCGCGCTCTGCTGCA 2221  
 2327 ACAGAAATGACAAACAACTGTGTATAGCTTCCCTTTCAAGGACATGACAAATTTTCAGG 2386  
 2222 GCAGAAATGACAGAGATGTCTATAGCTTGCCCTTCAAGACATGACAAATTTTCAGG 2281  
 2387 TTTGTTTTCTGCTCTAGACATCATTTCAAAATTTGGGTGTCAATTTCTTATGCTGTTCCAT 2446  
 2282 CTGTGTTTCTGCTCTAGACATCATTTCAAACTGGGTGTAAATTTCTTATGCTGTTCCAT 2341  
 2447 GACGACTTTGGAAGAGCTATTTTAAAGCTAGAAGTTGAAGCAGAAATTCACCAAGCAGA 2506  
 2342 GACGACATTTGGAAGAGCTATTTTAAAGCTAGAAGTTGAAGCAGAAATTCACCAAGCAGA 2401  
 2507 TTATAGTGTATTTACTAGACGCCACTGGAGGAAGAAATGGATTCAAATCTTTTGATGA 2566  
 2402 TTATAGTGTATTTACTAGACGCCACTGGAGGAAGAAATGGATTCAAAGTCTTTTGATGA 2461  
 2567 AATGGAACAGAGCTTACTATTTCTTCTGAAACCAAGCTTCTCTAGTGAGCACCAGAG 2626  
 2462 AATGGAACAGAGTTTACTATTTCTCTCTGAAACCAAGCTTCTCTAGTGAGCACCAGAG 2521  
 2627 CCTTTGGAAACAACAGATGTATACAATAGCAAGTTCATTTCTTTTACCTTTGAAACGTGA 2686  
 2522 TCTCTGGAAGACAGCAAGTATCTACAATAGCAAGTTCATTTCTCTCTCACTGGAAGCAGA 2581  
 2687 AAGTAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTT 2746  
 2582 AAGCAATCAGTGCGCTCTGTGTTGCTTCTGCTTTTAAATTTTTCAGTTCAGATTTT 2641  
 2747 TATGTTTTGTTGTTTCACTCTCTTTTAAATAATGCTGTGTTGCCATCAAACTGTGTCAGA 2806

2642 TATGTTTTGGTCCATCATTTCTTTAAAAATGCTGTGGTTCCCATCAAGCTGTTCCAGA 2701  
 2807 CTTATATTTTCTAAACCTGGAGACAACCAATATAAATAAATAAAGTCTGCTCTTCA 2866  
 2702 CTTGATTTTCTAAACCTGGAGATAAACCTCATATAAATAAATAAAGCTGCTGCTCA 2761  
 2867 AAATCTGCT-----GACTCAGATATCAGTGTATCTTATATAGCTTTTTTCA 2910  
 2762 AAATTTCTCTGCTGGAGCGGTGGAAGATTCAATATCAATGATCTTATGAGCTTTTTCA 2821  
 2911 CAAAGCAGAAATTAATGTTGACGATGATTAATGACAGTGAATATGATATCCGTTGGCTCC 2970  
 2822 CACAGCAGAAATTAATGTTGGCCATGTTTAAATGACAGTGAATATGATGCTGCTGCTC 2881  
 2971 ATAGTGGGCTTTAAATGTTGATGCAATCAGAAAAAGCACTATGTTTTCAGCTGTTTTCA 3030  
 2882 ATAGTGGGCTCTCAATGTTGGTGCAGTCAGAAAAAGCACTATGTTTACTGCTGTTTTCA 2941  
 3031 ACAGTACTATGTTTATTTCTTTTACCTATATTAATGATGAATATCAATAGTAACCTACTATCT 3090  
 2942 ACAGTACTATGTTTATTTCTTTTCCAGTAATGATGAACATCATTTAGTAACCTACTATCT 3001  
 3091 ATCATTTAAATGTTGACTGAACCATCCAGATCTGAGTACCCCATTTCTTTCAAGAAATTA 3150  
 3002 ATCATTTAAATGTTGACTGACACTATCCAGATCTGAGTACCCCGTTCTTCAAGAAATTA 3061  
 3151 CTGATATAGTTTTTAAAAATGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATGTTA 3210  
 3062 CTGACATTTGTTTTTAAAGTTGAGCTATATTTTCAAGCAGCTTTGCTTGGAAATCATGTTA 3121  
 3211 CTGCAATGCCACTTTACCTTTGCCATGGAATAATGACAGAAATCATAGATCAAGCTTATA 3270  
 3122 CTGCAATGCCACTTTACCTTTGCCATGGAATAATGACAGAAATCATAGATCAAGCTTATA 3181  
 3271 CTCAACTTAAATTTTCAAGCTCTTTTCCCATCTGCAATTTGGATTTGCAAGCTGTTGTG 3330  
 3182 CTCAGCTTAAATTTTCAAGCTCTTTTCCCATCTGCAATTTGGATTTGCAAGCTGTTGTG 3241  
 3331 ATATCCCTTATTTTATCATTTTATTTGATGCTAGGAGCTTATGCAATTTCAAT 3390  
 3242 ACATTTCCCTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3301  
 3391 ATGGATATATTTTATATCTGTTAAAGTTTCTTGTGCTGTTGTTGTTGTTGTTGTTGTTG 3450  
 3302 ACGGACTGTATTTTATCTCTGTTAAAGTTTCTTGTGCTGTTGTTTGTGCTCATTTGTT 3361  
 3451 TTCCATCAGTATTTCTGTTCACTTATATGCTTTTCACTTCACTTTTCAAGAAATTTTAA 3510  
 3362 TGCCCTCGGTTATTTCTGTTCACTTATATAGCTTGTGTTCACTTTTAAAGAAATTTTAA 3421  
 3511 CCAAGAAATTTTGGTCAATTTATCTATCTGTGACAGCATTTGGCTTGTGTCAGTCACTG 3570  
 3422 CCAAGAAATTTTGGTCAATTTATCTATCTGTGACAGCATTTGGCTTGTGTCAGTCACTG 3481  
 3571 AAATAACTTTTCTTATGGATACAAATGCAATTTCTTCTTCAATATGCTTTTGTATCA 3630  
 3482 AAATAACTTTTCTTCTGGGTTATGGAGTTACGGCTGTTTCTTCAATACCTTCTGCACTG 3541  
 3631 TCATTTCCAAATCTCATCCACTTCTAGTGTGCTGTTGTTTCTTCAATAAGATTTTCTTGA 3690  
 3542 CCATCCCAATCTACCCGCTCTCGGTTGTGTTGTTTCTTCAATAAGGTTTCTTGAAGA 3601  
 3691 ATGTACGAAAAAATGTGGACACCTATAATCAATGGGATAGGCTTTCACTAGCTGTTATAT 3750  
 3602 ATATACCAAAAACTGAGAAATGCTTCAATCCCTGGGACAGACTTTTGTGCTGTTAAATCA 3661  
 3751 CGCCCTTACCTGAGGTGTGCTGTTGTTTCTTCTTCAATATCTATGAGAAAAAATATG 3810  
 3662 TGCCCTTACCTGAGGTGTGCTGTTGTTTCTTCTTCAATATCTATGAGAAAAAATATG 3721  
 3811 GAGGAGATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAAGTCTAAAAATA 3870  
 3722 GAGGAGATCAATAAGAAAAAGATCCCTTTTTCAGGCGCTCTCTCAAAAAAGGCAACATA 3781

```
QY 3871 GGAAGCTTCAGAACACACAGACAATGAGATGAAGATGAAGATGTCARAGCTGAAGAC 3930
Db 3782 AGAAGTTTCAGAACACACACCCATCAATGAGATGAAGATGAAGATGTCARAGCTGAAGAC 3841
QY 3931 TAAAGTCAAGAGCTGATGGGTGCGAGTGTGTGAGGAGAAACCATCATATTATGGTCA 3990
Db 3842 TGAAGGTTAAAGAGCTGATGGGTGCGAGTGTGTGAGGAGAAACCGGCCATTATGGTAT 3901
QY 3991 GCAATTGCTAAGATATGATGACAAGAGATTTCTTCTTCAAGAAAAGTAAAGA 4050
Db 3902 ACAATTGCTAAGAAATATGATGACAAGAAATTTCTTCTTCAAGAAAAGTAAAGA 3961
QY 4051 AAGTGGCAACTAATACATCTCTTCTGTGTGAAAAAGGAGAGATCTTAGCACTATTGG 4110
Db 3962 AAGTAGCAACAATACCTCTCTTCTGTGTGAAAAAGGAGAGATCTTAGCACTATTGG 4021
QY 4111 GTCCAAATGGTGTGCAAAAAGCACAATTAATTAATTTCTGGTGGTGAATGAAACAA 4170
Db 4022 GTCCAAATGGTGTGCAAAAAGCAGCATTAATTAATTTCTGGTGGTGAATGAAACAA 4081
QY 4171 CTTGAGGCCAGGTATTTTAGGAGATTTCTTCTGAGACAAGTGAAGATGATGATCA 4230
Db 4082 CTTGAGGCCAGATTTCTTAGGAGATTTCTTCTGAGACAAGTGAAGATGATGATCA 4141
QY 4231 TGAAGTGTATGGTTACTGTCTCAGATAAAACCTTTGTGGCCAGATTAATGATGCA 4290
Db 4142 CCAAGTGTATGGTTACTGTCTCAGACAACCACTCTGGCCAGATTAATGATGCA 4201
QY 4291 AACATTTGAAATTTATGAGCTGTCAAGAGATGATGCAAGTGAAGATGATGCA 4350
Db 4202 AACATTTGAAATTTATGAGCTGTCAAGAGATGATGCAAGTGAAGATGATGCA 4261
QY 4351 TAAGTCGAATAACACATGCACTTATTTAAAGAAACATCTTCAGAAAGTGTAAAGAA 4410
Db 4262 TCAGTCGAATAACAAAGCCCTTGTATTTGAAAGAAACATCTTCAGAAAGTGTAAAGAA 4321
QY 4411 TACCTCAGGAATCAACAGAAAGTGTGTGTGCTTAAGTATGCTAGGATCTCAGA 4470
Db 4322 TCCTCAGGAGATCAAGCGAAGCTGTGTGTGCTCTAGTATGCTGGGAAATCTCAGG 4381
QY 4471 TTACTTTGCTAGATGAACCATCTACAGGTATGATGCCAAAGCCCAACAGACATGTGGC 4530
Db 4382 TGACACTGTGTGATGAGCGTCTACAGGTATGAGCCCAAGAGCCCAACAGACATGTGGC 4441
QY 4531 GACCAATTCGACTGATTAACATTAACAGAAAGAGGCTGCTATTCTGACCACTCACTATA 4590
Db 4442 GAGCCATTGCACTGCGTTTAAACACAGAAAGCGGCTGCCCTCTCTCACTAGCCATTACA 4501
QY 4591 TGGAGGAGGCGAGAGGCTGTGTGATCGAGTAGCTATCATGCTGTCTGGGCACTTAAGAT 4650
Db 4502 TGGAGAGGCGGAGGCTGTGTGACAGAGGTGGCCATCATGTGTCTCGGGCAGCTAAGAT 4561
QY 4651 GTATCGGAACAGTACAACTCTAAGAGTAAATTTGGAAAGAGGCTACTTTTGGAAATTA 4710
Db 4562 GCATTGGAACAGTACAACTCTAAGAGCAAGTTTGGCAAGAGGCTACTTTTGGAAATTA 4621
QY 4711 AATTGAAGACTCGATAGAAACCTAGAAAGTAGACCGCTTCAAGAGAAATTCAGTATA 4770
Db 4622 AACTAAGAGACTCGATAGAAACCTAGAAATAGATGCGCTTCAAGAGAAATTCAGTATA 4681
QY 4771 TTTTCCCAATGCAAGCCGCTCAGGAAGTTTTTCTTCTATTATTGCTTATATAATTCCTA 4830
Db 4682 TTTTCCCAATGCAAGCCGCTCAGGAAGTTTTTCTTCTATTATTGCTTATATAATTCCTA 4741
QY 4831 AGGAAGATCTCAGTCCCTTTTCAATCTTTTAAAGCTGGAAGAGCTTAACATGCTT 4890
Db 4742 AAGAAATGTCAGTCTCTCTCAGCTCTCTTGTGAAGTGGAGGAAGCAACACACTT 4801
QY 4891 TTGCCATTGAAGAATATAGCTTTTCTCAGCAACATTTGAAACAGCTTTTGTAGAACTCA 4950
Db 4802 TTGCCATTGAAGAATATAGCTTTTCTCAGCAACCTTGAACAGCTTTTGTAGAACTTA 4861
```

```
QY 4951 CTAAGAACACAGGAGGAGAGATTAATAGTTGTGGAACCTTTAAACACAGACACTTTGGTGG 5010
Db 4862 CTAAGGAGGAGGAGGAGAGATTAATAGCTGCGGGAACCTTTAAACACAGCAGCTTGGTGG 4921
QY 5011 AACGAACACACAGAGATAGAGTAG 5034
Db 4922 AAAGAAGGAGGAGAGACAGAGTCG 4945

RESULT 15
HSM806823
LOCUS Homo sapiens mRNA; cDNA DKFZp686K07118 (from clone DKFZp686K07118).
DEFINITION BX640746
ACCESSION BX640746
VERSION BX640746.1 GI:34364843
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 4078)
AUTHORS Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H.W., Weil, B., Amid, C., Oesinger, A., Fobo, G., Han, M. and
Wiemann, S.
COMMENT The German Human cDNA Consortium
DIRECT SUBMISSION
TITLE Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
JOURNAL Neuherberg, GERMANY
Cloned from S. Wilmann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686K07118) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
LOCATION/Qualifiers
1. 4078
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="17q24.3"
/clone="DKFZp686K07118"
/tissue_type="human fetal kidney"
/clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="fetal"
1. 4078
/genes="DKFZp686K07118"
<1..1295
/genes="DKFZp686K07118"
/notes="AP- binding cassette A5, N-terminus truncated,
differentially spliced"
/codon_start=1
/product="hypothetical protein"
/protein_id="CA245856.1"
/db_xref="GI:34364844"
/translation="VFVFVTFVAFPTGLMILLIESPKSLVWLPSPFCHCTFVIGIAQ
VMELEDFNFGASFNLTAGPYPLIITIMLTNSIFVLLAVLDQVTFGEFGRSS
LYLPKPSYMSKSNRYEELSNGVNGNIISFIEFVSFEFVKGKAIRISGIGKTRK
KGNVEALRNLSFDIYEGQITALLHSGTGKSTLNNILGCLCPSPGFSIYGHRYSE
IDSMFARKMIGICPOLDIHFVLTVEENLSILASIKGIPANNIIQEVQKVLDDMQ
TIKDQAKLSGSKREKLSIGLAVLGNPKILLDEPTAGMDPCSRHIVNLLKRXAN
RVTVFSTHWDENDILLADKANTVITSGMLKGVSSMFLKSKWGIYRLSMYIDKYCATE
SLSLVYKQHPGATLLQQNDQQLVSLPFDMDKFSGIA"
4041..4046
/genes="DKFZp686K07118"
4056
/genes="DKFZp686K07118"

polyA_signal
polyA_site

ORIGIN
```

Query Match		69.6%; Score 3812.6; DB 9; Length 4078;
Best Local Similarity		95.9%; Pred. No. 0;
Matches 4014; Conservative		0; Mismatches 4; Indels 169; Gaps 3;
QY	1101	GTGTAATTTTGTGTAAGTGGCTTTTGGATTTATGACCTTATGATTAATCTCTATAGAA 1160
DB	1	GTGTAATTTTGTGTAAGTGGCTTTTGGATTTATGACCTTATGATTAATCTCTATAGAA 60
QY	1161	AGTTTTCCAAATCGTTAGTGGCTTTTCACTCCCTTCTGTCACCTGACTTTTGTGATT 1220
DB	61	AGTTTTCCAAATCGTTAGTGGCTTTTCACTCCCTTCTGTCACCTGACTTTTGTGATT 120
QY	1221	GGTATTGACAGCTCATGATTTAGAGATTTTAAAGAGTGGCTTCAITTTCAAATTTG 1280
DB	121	GGTATTGACAGCTCATGATTTAGAGATTTTAAAGAGTGGCTTCAITTTCAAATTTG 180
QY	1281	ACTGCAGGCCATATCTCTAAATTTACAAATATCATCTCACACTTAAATAGTATATTC 1340
DB	181	ACTGCAGGCCATATCTCTAAATTTACAAATATCATCTCACACTTAAATAGTATATTC 240
QY	1341	TATGTCCTCTGGCTGTCTATCTTGATCAAGTCAATCCAGGGGAATTTGGCTTACGGAGA 1400
DB	241	TATGTCCTCTGGCTGTCTATCTTGATCAAGTCAATCCAGGGGAATTTGGCTTACGGAGA 300
QY	1401	TCATCTTTATATTTCTGAAGCTTCTATTTGGTCAAGAGCAAAAGAAATTTATGAGAG 1460
DB	301	TCATCTTTATATTTCTGAAGCTTCTATTTGGTCAAGAGCAAAAGAAATTTATGAGAG 360
QY	1461	TTATCAGAGGCAATGTTAATGGAATATTTAGTTTTAGTGAATATTTAGGCCATTTCT 1520
DB	361	TTATCAGAGGCAATGTTAATGGAATATTTAGTTTTAGTGAATATTTAGGCCATTTCT 420
QY	1521	TCAGAAATTTGTAGGAAGAGCCATTAAGATTTAGTGGTATTCAGAGACATACAGAAAG 1580
DB	421	TCAGAAATTTGTAGGAAGAGCCATTAAGATTTAGTGGTATTCAGAGACATACAGAAAG 480
QY	1581	AAGGGTGAATTTGGAGCTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAGATT 1640
DB	481	AAGGGTGAATTTGGAGCTTTGAGAAATTTGTCAATTTGACATATATGAGGGTCAGATT 540
QY	1641	ACTGCCCTTACTGGCCACAGTGGAGCAAGAGATGACATTCATGAATTTCTTTGTGGA 1700
DB	541	ACTGCCCTTACTGGCCACAGTGGAGCAAGAGATGACATTCATGAATTTCTTTGTGGA 600
QY	1701	CTCTGCCACCTTCTGATGGGTTTGATCTATATATGACACAGAGTCTCAGAAATAGAT 1760
DB	601	CTCTGCCACCTTCTGATGGGTTTGATCTATATATGACACAGAGTCTCAGAAATAGAT 660
QY	1761	GAAATGTTGAAGCAAGAAATGATTTGGCATTTGTCATCTATATGACACAGAGTCTCAGAAATAGAT 1820
DB	661	GAAATGTTGAAGCAAGAAATGATTTGGCATTTGTCATCTATATGACACAGAGTCTCAGAAATAGAT 720
QY	1821	GTTTTTCACAGTGAAGAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAAC 1880
DB	721	GTTTTTCACAGTGAAGAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAAC 780
QY	1881	AATATAATCAAGAGTGGAGAGGTTTACTAGATTTAGACATGACAGATCAAGAT 1940
DB	781	AATATAATCAAGAGTGGAGAGGTTTACTAGATTTAGACATGACAGATCAAGAT 840
QY	1941	AACCAAGCTTAAATTTAGTGGTGGTCAAAAGAGAGCTCTCATTTAGGAATTCCTGTT 2000
DB	841	AACCAAGCTTAAATTTAGTGGTGGTCAAAAGAGAGCTCTCATTTAGGAATTCCTGTT 900
QY	2001	CTTTGGAAACCAAGATACCTGCTGATGAAACCAAGCTGGAAATGGACCCCTGTTCT 2060
DB	901	CTTTGGAAACCAAGATACCTGCTGATGAAACCAAGCTGGAAATGGACCCCTGTTCT 960
QY	2061	CGACATATTTGTAATGGAATCTTTTAAATACAGAAAGCAATCGGCTGACAGTGTTCAGT 2120
DB	961	CGACATATTTGTAATGGAATCTTTTAAATACAGAAAGCAATCGGCTGACAGTGTTCAGT 1020
QY	2121	ACTCATTTTCATGATGAAGCTGACATTTCTTTGAGATGAGAAAGCTGTGATATCAAGGA 2180
DB		
DB	1021	ACTCATTTTCATGATGAAGCTGACATTTCTTGAGATAGAAAGCTGTGATATCAAGGA 1080
QY	2181	ATGCTGAAATGCTGTTGGTCTTCAATGTTCCCTCAAAAGTAATAGGGGATCGGCTACCGC 2240
DB	1081	ATGCTGAAATGCTGTTGGTCTTCAATGTTCCCTCAAAAGTAATAGGGGATCGGCTACCGC 1140
QY	2241	CTGAGCATGTACATAGACAAATATTTGGCCACAGAAATCTCTTTCTTCACTGGTTAAACAA 2300
DB	1141	CTGAGCATGTACATAGACAAATATTTGGCCACAGAAATCTCTTTCTTCACTGGTTAAACAA 1200
QY	2301	CATATACCTGGAGCTATTTTATACACAGATGACCAACACTTGTGTATAGCTTGCCT 2360
DB	1201	CATATACCTGGAGCTATTTTATACACAGATGACCAACACTTGTGTATAGCTTGCCT 1260
QY	2361	TTCAAGGACATGGACAAATTTTTCAG-----GT 2387
DB	1261	TTCAAGGACATGGACAAATTTTTCAGGAAATTTCTGATAGACAAAGGATGTTATTGATGT 1320
QY	2388	TTGTTTCTGCCCCCTAGACAGCTCATTTCAAATTTGGGTGTCTATTTCTTTATGGTGTTCATG 2447
DB	1321	TTGTTTCTGCCCCCTAGACAGCTCATTTCAAATTTGGGTGTCTATTTCTTTATGGTGTTCATG 1380
QY	2448	ACGACTTTTGAAGACGCTATTTTAAAGCTAGAAATTTGAAGCAGAAATTTGACCAAGCAGAT 2507
DB	1381	ACGACTTTTGAAGACGCTATTTTAAAGCTAGAAATTTGAAGCAGAAATTTGACCAAGCAGAT 1440
QY	2508	TATAGTGTATTTACTCAGAGCCACTGAGAGAGAAATGGATTCAAAATCTTTTGTATGAA 2567
DB	1441	TATAGTGTATTTACTCAGAGCCACTGAGAGAGAAATGGATTCAAAATCTTTTGTATGAA 1500
QY	2568	ATGGAACAGAGCTTACTTATTTCTGTAACCAAGGCTTCTCTAGTGAGCACCATGAGC 2627
DB	1501	ATGGAACAGAGCTTACTTATTTCTGTAACCAAGGCTTCTCTAGTGAGCACCATGAGC 1560
QY	2628	CTTTGGAAACACAGATGATACATAGCAAGTTTCATTTCTTTTACCTTGAACCTGAA 2687
DB	1561	CTTTGGAAACACAGATGATACATAGCAAGTTTCATTTCTTTTACCTTGAACCTGAA 1620
QY	2688	AGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTT 2747
DB	1621	AGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTT 1680
QY	2748	ATGTTTTTGGTTTCACTCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTTCAGAC 2807
DB	1681	ATGTTTTTGGTTTCACTCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTTCAGAC 1740
QY	2808	TTATATTTCTTAAACCTGGAGACAAACCAATAAATAAATAAATAAATAAATAAATAAATAA 2867
DB	1741	TTATATTTCTTAAACCTGGAGACAAACCAATAAATAAATAAATAAATAAATAAATAAATAA 1800
QY	2868	AAATCTGCT-----GACTCAGATATCAGTGATCTTTATTAGCTTTTTCAC 2911
DB	1801	AAATCTGCTGCTGAGAGTGTGGAAGACTCAGATATCAGTGATCTTTATTAGCTTTTTCAC 1860
QY	2912	AAGCCAGAACATTAATGGTGACGATTAATAGACAGTACTATGATCGTGCTCCCCA 2971
DB	1861	AAGCCAGAACATTAATGGTGACGATTAATAGACAGTACTATGATCGTGCTCCCCA 1920
QY	2972	TAGTGGCGCTTTAAATGATGATCAITTCAGAAAAAGGACTATGTTTTTTCAGCTGTTTTCAA 3031
DB	1921	TAGTGGCGCTTTAAATGATGATCAITTCAGAAAAAGGACTATGTTTTTTCAGCTGTTTTCAA 1980
QY	3032	CAGTACTATGGTTTATTTCTTTACTATATAGTGAATATCATTAGTAACTACTATCTTTA 3091
DB	1981	CAGTACTATGGTTTATTTCTTTACTATATAGTGAATATCATTAGTAACTACTATCTTTA 2040
QY	3092	TCATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCAATCTTTTCAAGAAATAC 3151
DB	2041	TCATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCAATCTTTTCAAGAAATAC 2100
QY	3152	TGATATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTTCTTGGATCATTTGTAC 3211
DB		

Db	2101	TGATATAGTTTTTAAAATTGAGCTGTATTTTCAAGCAGAGCTTGGCTTGGAAATCAATTGTTAC	2161
Qy	3212	TGCAATGCACCTTACTTTTGGCATGGAAAATGCGAGAGAATCATAAAGATCAAAAGCTTATAC	3271
Db	2161	TGCAATGCACCTTACTTTTGGCATGGAAAATGCGAGAGATCATPAAGATCAAAAGCTTATAC	2220
Qy	3272	TCAACTTAAACTTTTCAGGTCTTTTGGCCATCTGGATATATGGATTGGACAAGCTGTGTGTGA	3331
Db	2221	TCAACTTAAACTTTTCAGGTCTTTTGGCCATCTGGCATATATGGATTGGACAAGCTGTGTGTGA	2280
Qy	3332	TATCCCCTTATTTTTTTATCAATCTTATTTTGTATGCTAGGAAGCTTATTTGGCATTTCAATTA	3391
Db	2281	TATCCCCTTATTTTTTTATCAATCTTATTTTGTATGCTAGGAAGCTTATTTGGCATTTCAATTA	2340
Qy	3392	TGGATTATATTTTTTATCTGTAAAGTTCCCTTTCGTGTGGTTTTTTTGGCCTTATTTGGTTATGT	3451
Db	2341	TGGATTATATTTTTTATCTGTAAAGTTCCCTTTCGTGTGGTTTTTTTGGCCTTATTTGGTTATGT	2400
Qy	3452	TCCATCAGTTATTTCTGTTCACCTATATATGCTTTTCCACCTTAAAGAAAATTTTAAATAC	3511
Db	2401	TCCATCAGTTATTTCTGTTCACCTATATATGCTTTTCCACCTTAAAGAAAATTTTAAATAC	2460
Qy	3512	CAAGAATTTTGGTCATTTATCTATTTCTGTGGCAGCGTTGGCTTGTATTTGCAATCACTGA	3571
Db	2461	CAAGAATTTTGGTCATTTATCTATTTCTGTGGCAGCGTTGGCTTGTATTTGCAATCACTGA	2520
Qy	3572	AATAAATTTCTTTATGGGATACACAATTTGCGAATTTCTTCAATATGCCCTTTTGTATCAT	3631
Db	2521	AATAAATTTCTTTATGGGATACACAATTTGCGAATTTCTTCAATATGCCCTTTTGTATCAT	2580
Qy	3632	CATTCCAAATCTATCCACTTCTAGTGTGCCGTGATTTCTTTTCATAAAGATTTCTTTGGAAGAA	3691
Db	2581	CATTCCAAATCTATCCACTTCTAGTGTGCCGTGATTTCTTTTCATAAAGATTTCTTTGGAAGAA	2640
Qy	3692	TGTACGAAAAAATGTGGACACCTATATATCCATGGGATAGGCTTTTCACTAGCTGTTATATC	3751
Db	2641	TGTACGAAAAAATGTGGACACCTATATATCCATGGGATAGGCTTTTCACTAGCTGTTATATC	2700
Qy	3752	GCCTTACCTCGAGTGTACTGTGGATTTTTCCTCTTACAATACTATGAGAAAAAATATGG	3811
Db	2701	GCCTTACCTCGAGTGTACTGTGGATTTTTCCTCTTACAATACTATGAGAAAAAATATGG	2760
Qy	3812	AGGCAGATCAATTAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAAGCTTAAAAATAG	3871
Db	2761	AGGCAGATCAATTAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAAGCTTAAAAATAG	2820
Qy	3872	GAAGCTTCCAGAACCCACGACATGAGGATGAAGATGTCAAGAGTCTCAAGCTTGAAGACT	3931
Db	2821	GAAGCTTCCAGAACCCACGACATGAGGATGAAGATGTCAAGAGTCTCAAGCTTGAAGACT	2880
Qy	3932	AAAGGTCAAAAGAGCTGATGGGTTGCCAGTGTCTGTGAGGAGAAACCATCCATTATGTCTAG	3991
Db	2881	AAAGGTCAAAAGAGCTGATGGGTTGCCAGTGTCTGTGAGGAGAAACCATCCATTATGTCTAG	2940
Qy	3992	CAATTTGCATAAAGAATATGATCAAGAAGAATTTTCTTTTCAAGAAAAAGTAAGAA	4051
Db	2941	CAATTTGCATAAAGAATATGATCAAGAAGAATTTTCTTTTCAAGAAAAAGTAAGAA	3000
Qy	4052	AGTGGCAACTAAATACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGG	4111
Db	3001	AGTGGCAACTAAATACATCTCTTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGG	3060
Qy	4112	TCCAAATGGTGTGGCAAAAAGCAATTTATTAATATTTCTGGTTGGTGTGATTTAGAACCAAC	4171
Db	3061	TCCAAATGGTGTGGCAAAAAGCAATTTATTAATATTTCTGGTTGGTGTGATTTAGAACCAAC	3120
Qy	4172	TTCAGGCCAGGTATTTTTTAGAGATTTCTTCCAGAGACAAGTGAAGATGATGATTCAC	4231
Db	3121	TTCAGGCCAGGTATTTTTTAGAGATTTCTTCCAGAGACAAGTGAAGATGATGATTCAC	3180
Qy	4232	GAAGTGATGGGTTACTGTCTCAGATAAACCCCTTTTGGCCAGATACTACATTCAGGA	4291
Db	3181	GAAGTGATGGGTTACTGTCTCAGATAAACCCCTTTTGGCCAGATACTACATTCAGGA	3240

QY	4292	ACATTTTGAATTTATCGAGCTGTCAAAGAAATGAGTGCAGATGAAGAAGTCAT	4351
DB	3241	ACATTTTGAATTTATCGAGCTGTCAAAGAAATGAGTGCAGATGAAGAAGTCAT	3300
QY	4352	AAGTCGAATAACACATGCATCTGATTTAAAGAAACATCTTCAGAAAGACTGTAAAGAAACT	4411
DB	3301	AAGTCGAATAACACATGCATCTGATTTAAAGAAACATCTTCAGAAAGACTGTAAAGAAACT	3360
QY	4412	ACCTGCAGGAATCAAACGAAAGTTGTGTTTTGCTCTTAAGTATGCTTAGGGAATCCTCAGAT	4471
DB	3361	ACCTGCAGGAATCAAACGAAAGTTGTGTTTTGCTCTTAAGTATGCTTAGGGAATCCTCAGAT	3420
QY	4472	TACTTTGCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAAACAGACACATGTGGCG	4531
DB	3421	TACTTTGCTAGATGAACCATCTACAGGTATGGATCCCAAGCCAAAACAGACACATGT	3476
QY	4532	AGCAATTCGAATGTCATTTAAAACAGAAAGCGGGCTGTATTTCTGACCACTCACTATAT	4591
DB	3477	-----	3476
QY	4592	GGAGGAGCAGAGGCTGCTGTGATCGAGTAGCTATCATGGTGTCTGGCGAGTTAAGATG	4651
DB	3477	-----GATG	3480
QY	4652	TATCGGAACAGTCAACATCTAAAGAGTAAATTTGAAAAAGGTACTTTTGGAAATTA	4711
DB	3481	TATCGGAACAGTCAACATCTAAAGAGTAAATTTGAAAAAGGTACTTTTGGAAATTA	3540
QY	4712	ATTGAAGACTGGATAGAAAACCTAGAAGTAGACCCCTTCAAGAGAAATTCAGATATAT	4771
DB	3541	ATTGAAGACTGGATAGAAAACCTAGAAGTAGACCCCTTCAAGAGAAATTCAGATATAT	3600
QY	4772	TTTCCCAATGCAAGCGGTGAGAAAAGTTTTCTTCTATTTGGCTTATAAAATTCCTAA	4831
DB	3601	TTTCCCAATGCAAGCGGTGAGAAAAGTTTTCTTCTATTTGGCTTATAAAATTCCTAA	3660
QY	4832	GGAGATGTTTCAGTCCCTTTTCAAACTCTTTTAAAGCTGGAAGAGCTTAAACATGCTTT	4891
DB	3661	GGAGATGTTTCAGTCCCTTTTCAAACTCTTTTAAAGCTGGAAGAGCTTAAACATGCTTT	3720
QY	4892	TGCCATTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAAACAGGTTTTGTAGAAGCTCAC	4951
DB	3721	TGCCATTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAAACAGGTTTTGTAGAAGCTCAC	3780
QY	4952	TAAAGAACAGAGGAGAGATTAATAGTTGTGGAACCTTTAAACAGACACACTTTGGTGGGA	5011
DB	3781	TAAAGAACAGAGGAGAGATTAATAGTTGTGGAACCTTTAAACAGACACACTTTGGTGGGA	3840
QY	5012	ACGAAACAAGAGATAGAGTAGTATTTTGAAATTTGTTATTTGGTCTGCTTACTGGGA	5071
DB	3841	ACGAAACAAGAGATAGAGTAGTATTTTGAAATTTGTTATTTGGTCTGCTTACTGGGA	3900
QY	5072	CTTCTCTTTTTCATTAATTTTAACCTTTGGTTTTAAAAAGTTTTTTTTATTGGAAATGGTAA	5131
DB	3901	CTTCTCTTTTTCATTAATTTTAACCTTTGGTTTTAAAAAGTTTTTTTTATTGGAAATGGTAA	3960
QY	5132	CTGAGAAACAAGAACGCACTTGAAATTTTTCTPAAGCTCCTTAATTAAGAAATGCTGTGGTT	5191
DB	3961	CTGAGAAACAAGAACGCACTTGAAATTTTTCTPAAGCTCCTTAATTAAGAAATGCTGTGGTT	4020
QY	5192	GTGTGTTTTGCTTTTCTTTAAATAAACAAGTATGATATAATTAAGTGAA	5238
DB	4021	GTGTGTTTTGCTTTTCTTTAAATAAACAAGTATGATATAATTAAGTGAA	4067

Search completed: April 12, 2004, 05:29:01  
Job time : 13683 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 11, 2004, 21:13:24 ; Search time 1299 Seconds  
(without alignments)  
17905.223 Million cell updates/sec

Title: US-10-090-458-4

Perfect score: 5475

Sequence: 1 gcgcgcgcgcctcgacacag.....cataagcaatggaagtt 5475

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001as.\*

5: Geneseq2001bs.\*

6: Geneseq2002s.\*

7: Geneseq2003as.\*

8: Geneseq2003bs.\*

9: Geneseq2003cs.\*

10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5474	100.0	5475	7	ABNS7751 cDNA enco
2	5373.2	98.1	6523	6	ABNS9594 Human ATP
3	5356.6	97.8	5463	7	ABNS7749
4	5135.6	93.8	6369	6	AAD33648 Human TRI
5	5014.6	91.6	5262	6	AAD37620 Human tra
6	4922.6	89.9	4929	6	AAD37618 Human tra
7	4811.6	87.9	4917	7	ABNS7750 Coding se
8	4759	86.9	4785	6	AAD37619 Human tra
9	3175.8	58.0	3268	7	ADA33416
10	2683.2	49.0	2723	9	ADCS1606 Human cod
11	1791	32.7	3928	6	ABK35707 cDNA sequ
12	1546.4	28.2	1548	6	ABK35706 cDNA sequ
13	1400	25.6	2481	7	ABZ35938 Human sec
14	1182	21.6	1346	6	ABL64459 Stomach c
15	1182	21.6	1346	6	ABL63763 Breast ca
16	1182	21.6	1346	6	ABNS6909 Gene #340
17	1181.6	21.6	1506	3	AAC81717 Human sec
18	1149.8	21.0	6181	6	ABNS9597 Human ATP
19	1149.2	21.0	5722	7	ABZ22923 Human ABC
20	1121.4	20.5	5981	6	ABNS9596 Human ATP
21	1119.8	20.5	4875	7	ABNS7519 Human ABC
22	1119.8	20.5	5018	7	ABNS7518 Human ABC
23	1031.4	19.9	5797	8	ADA20300 Human ATP

## ALIGNMENTS

## RESULT 1

ABNS7751  
ID ABNS7751 standard; cDNA; 5475 BP.

XX AC ABNS7751;

XX DT 04-FEB-2003 (first entry)

XX DE cDNA encoding novel human ATP binding cassette ABCA5 transporter #2.

XX KW Human; ATP binding cassette; ABC; ABCA5; transporter;

XX KW neurotoxin transport; beta-amyloid peptide; chromosome mapping;

XX KW blood brain barrier transport; tissue typing; predictive medicine;

XX KW ABCA5 mediated disorder; ABCA5 related disorder; gene therapy; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 114..5042

XX FT /\*tag= a

XX FT /product= "ABCA5"

XX FT /note= "ATP binding cassette (ABC) A5 transporter"

XX US2002123107-A1.

XX PN 05-SEP-2002.

XX PD 01-MAR-2002; 2002US-00090458.

XX PF 02-MAR-2001; 2001US-0272885P.

XX PR (ACTI-) ACTIVE PASS PHARM INC.

XX PA Chen H, Kilinski L, Le Bihan S;

XX PI WPI; 2003-066798/06.

XX DR P-PSDB; ABG72424.

XX XX Novel isolated ATP binding cassette transporter family polypeptide,

XX PT ABCA5, useful for treating disorders associated with aberrant or unwanted

XX PT ABCA5 transporter expression or activity.

XX PS Claim 2; Page 39-42; 52pp; English.

XX CC The invention describes an isolated ATP binding cassette (ABC)

XX CC transporter family polypeptide (1), designated ABCA5. (1) or the

ABNS9595 Human ATP  
Aad47363 Human tra  
Abq77736 Human ABC  
Abz22922 Human ABC  
Add29639 Mouse tum  
Aad33675 Human TRI  
Aad49503 Human TRI  
Aa144690 Human tra  
Abv99426 Human NOV  
Acc47571 Drug tran  
Abt10141 Human bre  
Aad33650 Human TRI  
Abv99425 Human NOV  
Aad33657 Human TRI  
Abz23526 Human pur  
Aas63176 Human sec  
Aas53722 Human cod  
Aas70746 DNA enco  
Aas66371 DNA enco  
Aak94093 Human ful  
Ab199789 Human sec  
Ab187826 Human ova



QY 1741 ACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATGGCAATTTGTCCAC 1800  
DB 1741 ACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATGGCAATTTGTCCAC 1800  
QY 1801 AGTTAGATATACACTTTGATGTTTTCACAGTAGAAGAAAAATTTCAATTTTGGCTTCAA 1860  
DB 1801 AGTTAGATATACACTTTGATGTTTTCACAGTAGAAGAAAAATTTCAATTTTGGCTTCAA 1860  
QY 1861 TCAGAGGATACAGCCCAACATATAATACAGAAAGTGCAGAGGTTTACTAGATTTAG 1920  
DB 1861 TCAGAGGATACAGCCCAACATATAATACAGAAAGTGCAGAGGTTTACTAGATTTAG 1920  
QY 1921 ACATGCAGACTATCAAGAGTACCAAGCTTAAAAATTAAGTGGTGGTCAAAAAGAGC 1980  
DB 1921 ACATGCAGACTATCAAGAGTACCAAGCTTAAAAATTAAGTGGTGGTCAAAAAGAGC 1980  
QY 1981 TGTCAATAGGAATGCTGTTCTTGGGACCCAAAGATACGTGCTAGATGAACCAACAG 2040  
DB 1981 TGTCAATAGGAATGCTGTTCTTGGGACCCAAAGATACGTGCTAGATGAACCAACAG 2040  
QY 2041 CTGGGAATGACCCCTGTTCTCGACATATTGTATGAATCTTTTAAAAATACAGAAAGCCA 2100  
DB 2041 CTGGGAATGACCCCTGTTCTCGACATATTGTATGAATCTTTTAAAAATACAGAAAGCCA 2100  
QY 2101 ATCGGGTGACAGTGTTCAGTACTCAATTCATGGATGAAGCTGACATTTCTGCAGATAGGA 2160  
DB 2101 ATCGGGTGACAGTGTTCAGTACTCAATTCATGGATGAAGCTGACATTTCTGCAGATAGGA 2160  
QY 2161 AAGCTGTGTGATACAGAGGAATGCTGAAATGTTGGTGTCTTCAATGTTCTCAAAAGTA 2220  
DB 2161 AAGCTGTGTGATACAGAGGAATGCTGAAATGTTGGTGTCTTCAATGTTCTCAAAAGTA 2220  
QY 2221 AATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACAGAAATCTC 2280  
DB 2221 AATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACAGAAATCTC 2280  
QY 2281 TTTCTTCACTGGTTAAAACACATATACCTGGAGCTACTTTATACACAGAAATGACCAAC 2340  
DB 2281 TTTCTTCACTGGTTAAAACACATATACCTGGAGCTACTTTATACACAGAAATGACCAAC 2340  
QY 2341 AACTTGTGTATAGTGTGCTTTCAAGGACATGGACAAATTTTCAGGTTGTTTCTGCCC 2400  
DB 2341 AACTTGTGTATAGTGTGCTTTCAAGGACATGGACAAATTTTCAGGTTGTTTCTGCCC 2400  
QY 2401 TAGACAGTCATTCAAATTTGGGTGTCAATTTCTTATAGTGTGTTTCCATGACGACTTTGGAAG 2460  
DB 2401 TAGACAGTCATTCAAATTTGGGTGTCAATTTCTTATAGTGTGTTTCCATGACGACTTTGGAAG 2460  
QY 2461 ACGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATGACCAAGCATTTATAGTGTATTTA 2520  
DB 2461 ACGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGAACAGCATTTATAGTGTATTTA 2520  
QY 2521 CTCAGCAGCCATCGGAGGAGAAATGGATTCAAAATCTTTTGAATAATGAAACAGAGCT 2580  
DB 2521 CTCAGCAGCCATCGGAGGAGAAATGGATTCAAAATCTTTTGAATAATGAAACAGAGCT 2580  
QY 2581 TACTTATTTCTGTAACCAAGGCTTCTAGTGCAGCACCAGCTTGGGAAACAAAC 2640  
DB 2581 TACTTATTTCTGTAACCAAGGCTTCTAGTGCAGCACCAGCTTGGGAAACAAAC 2640  
QY 2641 AGATGTATACATAGCAAAAGTTTCAATTTCTTTACCTTGAACCGTGAAGATGAATCAGTGA 2700  
DB 2641 AGATGTATACATAGCAAAAGTTTCAATTTCTTTACCTTGAACCGTGAAGATGAATCAGTGA 2700  
QY 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTTGGTTC 2760  
DB 2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTTGGTTC 2760  
QY 2761 ATCAGTCTTTTAAAAATGCTGGTTCCTCAATCAAACTGTTCCAGACTTATATTTCTAA 2820  
DB 2761 ATCAGTCTTTTAAAAATGCTGGTTCCTCAATCAAACTGTTCCAGACTTATATTTCTAA 2820

QY 2821 AACCTGGAGCAAAACACATAAATACAAAACAGAGTCTGCTTCTTCAAAATTCGTGACT 2880  
DB 2821 AACCTGGAGCAAAACACATAAATACAAAACAGAGTCTGCTTCTTCAAAATTCGTGACT 2880  
QY 2881 CAGATATCAGTGTATCTTATTAGCTTTTTCACAGCCAGAACATAATGGTGACAGATGATTA 2940  
DB 2881 CAGATATCAGTGTATCTTATTAGCTTTTTCACAGCCAGAACATAATGGTGACAGATGATTA 2940  
QY 2941 ATGACAGTGTATCTATCGTGGCTCCCATAGTGGGCTTTAAATGTAGATCAGTTCAG 3000  
DB 2941 ATGACAGTGTATCTATCGTGGCTCCCATAGTGGGCTTTAAATGTAGATCAGTTCAG 3000  
QY 3001 AAAAGGACTATGTTTGGAGCTGTTTTCACACAGTACTACTGTTTATTCTTTACCTATAT 3060  
DB 3001 AAAAGGACTATGTTTGGAGCTGTTTTCACACAGTACTACTGTTTATTCTTTACCTATAT 3060  
QY 3061 TAGTGAATATCATAGTAATCACTACTTATCAATTTAAATGTGACTGAACCATCAGA 3120  
DB 3061 TAGTGAATATCATAGTAATCACTACTTATCAATTTAAATGTGACTGAACCATCAGA 3120  
QY 3121 TCTGGAGTACCCCATCTTTCAGAAATTAATGATAGTATGATAGTATGATGATGATGAT 3180  
DB 3121 TCTGGAGTACCCCATCTTTCAGAAATTAATGATAGTATGATAGTATGATGATGATGAT 3180  
QY 3181 TTCAGCAGCTTTGCTTGGAAATCAITGTTTACTGCAATGGCACTTACTTTGCCATGAAA 3240  
DB 3181 TTCAGCAGCTTTGCTTGGAAATCAITGTTTACTGCAATGGCACTTACTTTGCCATGAAA 3240  
QY 3241 ATCAGAGAAATCAATAGATCAAAAGCTTACTCAACTTAACTTTCAGGCTCTTTGCCAT 3300  
DB 3241 ATCAGAGAAATCAATAGATCAAAAGCTTACTCAACTTAACTTTCAGGCTCTTTGCCAT 3300  
QY 3301 CTGCATATGGATTTGGACAAAGCTGTTTGTGATATCCCTTATTTTATATCATCTTATTT 3360  
DB 3301 CTGCATATGGATTTGGACAAAGCTGTTTGTGATATCCCTTATTTTATATCATCTTATTT 3360  
QY 3361 TGATGCTAGGAAGCTTATTTGGCAATTTTCAATTTAGGATTTATTTTATCTGTAAGTCTC 3420  
DB 3361 TGATGCTAGGAAGCTTATTTGGCAATTTTCAATTTAGGATTTATTTTATCTGTAAGTCTC 3420  
QY 3421 TTGCTGTGGTTTTTGGCTTATTTGTTTATGTTTCCATCAGTATTTCTGTTCACTTATATTG 3480  
DB 3421 TTGCTGTGGTTTTTGGCTTATTTGTTTATGTTTCCATCAGTATTTCTGTTCACTTATATTG 3480  
QY 3481 CTTCTTTTCACTTTAGAAAATTTTAAATACCAAGAAATTTGGTCAATTTCTATTTCTG 3540  
DB 3481 CTTCTTTTCACTTTAGAAAATTTTAAATACCAAGAAATTTGGTCAATTTCTATTTCTG 3540  
QY 3541 TGGCAGCGTTGGCTTGTATTTGCAATCACTGAAATAACTTTTCTTTATGGGATACACAATTG 3600  
DB 3541 TGGCAGCGTTGGCTTGTATTTGCAATCACTGAAATAACTTTTCTTTATGGGATACACAATTG 3600  
QY 3601 CAATCTATCTTCAATATGCTTTTGTATCATCTCAATCTATCTATCTATCTAGGTTGCC 3660  
DB 3601 CAATCTATCTTCAATATGCTTTTGTATCATCTCAATCTATCTATCTATCTAGGTTGCC 3660  
QY 3661 TGATTTCTTTTCAAAAAGATTTCTTGAAGAAATGACGAAAATAATGCGACACCTATAATC 3720  
DB 3661 TGATTTCTTTTCAAAAAGATTTCTTGAAGAAATGACGAAAATAATGCGACACCTATAATC 3720  
QY 3721 CATGGGATAGGCTTTTCAAGTGTGTTTATATGCGCTTATCTGAGTGTGTACTGTGATTT 3780  
DB 3721 CATGGGATAGGCTTTTCAAGTGTGTTTATATGCGCTTATCTGAGTGTGTACTGTGATTT 3780  
QY 3781 TCCTCTTCAATCAATAGAAAATAATGAGGAGAGATCAATAGAAAAGATCCCTTTT 3840  
DB 3781 TCCTCTTCAATCAATAGAAAATAATGAGGAGAGATCAATAGAAAAGATCCCTTTT 3840  
QY 3841 TCAGAAACCTTTTCAACGAAAGTCTTAAATAATAGGAAGCTTCCAGAACCAACAGCAATGAGG 3900  
DB 3841 TCAGAAACCTTTTCAACGAAAGTCTTAAATAATAGGAAGCTTCCAGAACCAACAGCAATGAGG 3900  
QY 3901 ATGAAAGTGAAGATGTCAAGCTGAAAGACTAAAGGTCAAGAGCTGATGGTGGTCCAGT 3960



Db 3901 ATGAAGTGAAGATGTCAAGCTCAAGAGCTAAAGGTCAGAGAGCTGATGGTGGCAGT 3960  
QY 3961 GTTGTGAGGAGAAACCATCCATTATGTGCAGCAATTTGCAATAAGAAATATGATGACAAGA 4020  
Db 3961 GTTGTGAGGAGAAACCATCCATTATGTGCAGCAATTTGCAATAAGAAATATGATGACAAGA 4020  
QY 4021 AAGATTTTCTTTCTTCAAGAAAAGTAAAGAAAGTGGCAACTTAATACATCTCTTCTGTG 4080  
Db 4021 AAGATTTTCTTTCTTCAAGAAAAGTAAAGAAAGTGGCAACTTAATACATCTCTTCTGTG 4080  
QY 4081 TGAAAAGAGGAGAGATCTTAGAGATTTTGGTCCAAATGTGGTGGCAAGAGCAACAATTA 4140  
Db 4081 TGAAAAGAGGAGAGATCTTAGAGATTTTGGTCCAAATGTGGTGGCAAGAGCAACAATTA 4140  
QY 4141 TTAATATTCTGGTGGTGATATTGAACCAACTTCAGGCCAGGTATTTTATAGGAGATTTT 4200  
Db 4141 TTAATATTCTGGTGGTGATATTGAACCAACTTCAGGCCAGGTATTTTATAGGAGATTTT 4200  
QY 4201 CTTGAGACAAAGTGAAGATGATGATTCATGGAAGTGTATGGGTACTGTCTCAGATTA 4260  
Db 4201 CTTGAGACAAAGTGAAGATGATGATTCATGGAAGTGTATGGGTACTGTCTCAGATTA 4260  
QY 4261 ACCCTTTGTGCCAGATACATTCAGGAAACATTTTGAATTTATGAGAGCTGTCAAAG 4320  
Db 4261 ACCCTTTGTGCCAGATACATTCAGGAAACATTTTGAATTTATGAGAGCTGTCAAAG 4320  
QY 4321 GAATGAGTGAAGTGAAGATGATGATTCATGGAAGTGTATGGGTACTGTCTCAGATTA 4380  
Db 4321 GAATGAGTGAAGTGAAGATGATGATTCATGGAAGTGTATGGGTACTGTCTCAGATTA 4380  
QY 4381 AAGACATCTTCAGAAAGTCTGTAAGAAATCTGAGGAACTCAAGCAAGTGTGTGTT 4440  
Db 4381 AAGACATCTTCAGAAAGTCTGTAAGAAATCTGAGGAACTCAAGCAAGTGTGTGTT 4440  
QY 4441 TTGCTCTAAGTATGTAGGAAATCTCAGATTAATTTGCTAGATGAACCATCTACAGTA 4500  
Db 4441 TTGCTCTAAGTATGTAGGAAATCTCAGATTAATTTGCTAGATGAACCATCTACAGTA 4500  
QY 4501 TGGATCCCAAGCCAAACAGACATGTGCGAGCAATTCGAATCTGCAATTTAAAAACAGAA 4560  
Db 4501 TGGATCCCAAGCCAAACAGACATGTGCGAGCAATTCGAATCTGCAATTTAAAAACAGAA 4560  
QY 4561 AGCGGGCTCTATTCTGACCACTACTATATGAGGAGGAGGAGCTGTCTGTGATCGAG 4620  
Db 4561 AGCGGGCTCTATTCTGACCACTACTATATGAGGAGGAGGAGCTGTCTGTGATCGAG 4620  
QY 4621 TAGCTATCATGTGTCTGGGAGTTAAGATGTATCGGAACAGTACCAATCTAAAGAGTA 4680  
Db 4621 TAGCTATCATGTGTCTGGGAGTTAAGATGTATCGGAACAGTACCAATCTAAAGAGTA 4680  
QY 4681 AATTTGGAAAAGGCTACTTTTGGAAATTAATTTGAAGAGCTGGATGAAAACCTAGAAG 4740  
Db 4681 AATTTGGAAAAGGCTACTTTTGGAAATTAATTTGAAGAGCTGGATGAAAACCTAGAAG 4740  
QY 4741 TAGACCGCTTCAAGAGAAATTCAGTATATTTTCCGAATTCGAAGCGCTCAGAAAGTT 4800  
Db 4741 TAGACCGCTTCAAGAGAAATTCAGTATATTTTCCGAATTCGAAGCGCTCAGAAAGTT 4800  
QY 4801 TTTCTCTATTATTTGGCTTATAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCACAATCTT 4860  
Db 4801 TTTCTCTATTATTTGGCTTATAAATTCCTAAGGAAGATGTTTCAGTCCCTTTTCACAATCTT 4860  
QY 4861 TTTTAAAGTGAAGAGTAAACATGCTTTTGCAATTTGAAGATATAGCTTTTCTCAAG 4920  
Db 4861 TTTTAAAGTGAAGAGTAAACATGCTTTTGCAATTTGAAGATATAGCTTTTCTCAAG 4920  
QY 4921 CACATTTGAACAGGTTTTTGTAGAACTCCTTAAAGAAACAGAGAGGAGAAATAGTT 4980  
Db 4921 CACATTTGAACAGGTTTTTGTAGAACTCCTTAAAGAAACAGAGAGGAGAAATAGTT 4980  
QY 4981 GTGGAATTTAAACAGCACTTTTGGTGGAAACGAAACCAAGAGATAGATGATTTT 5040

Db 4981 GTGGAATTTAAACAGCACTTTGGTGGAAACGAAACCAAGAGATAGATGATTTT 5040  
QY 5041 GAAATTTGATTTGTTGGTCTCTTACTCGGAGCTTCTTTTCTTTTCACTAATTTTAACTT 5100  
Db 5041 GAAATTTGATTTGTTGGTCTCTTACTCGGAGCTTCTTTTCTTTTCACTAATTTTAACTT 5100  
QY 5101 TGGTTTAAAGAGTTTATTTGGAATGTAATCTGAGAACCAAGAACGACCTTGAAATTT 5160  
Db 5101 TGGTTTAAAGAGTTTATTTGGAATGTAATCTGAGAACCAAGAACGACCTTGAAATTT 5160  
QY 5161 TTCTAAGTCTCTTAAATGAAATGCTGTGTTGTGTTTCTTTTCTTTTAAATTTAAACG 5220  
Db 5161 TTCTAAGTCTCTTAAATGAAATGCTGTGTTGTGTTTCTTTTCTTTTAAATTTAAACG 5220  
QY 5221 TATGATATTAATTAAGTGAAGTGCATGTTTGTATTAAGTATATTAAGTATATTAAGTAT 5280  
Db 5221 TATGATATTAATTAAGTGAAGTGCATGTTTGTATTAAGTATATTAAGTATATTAAGTAT 5280  
QY 5281 ATGTGATCTTTTTCACCAATCAGAAACAGTCTGCTGAAATTTGATTTAAAGAAATTT 5340  
Db 5281 ATGTGATCTTTTTCACCAATCAGAAACAGTCTGCTGAAATTTGATTTAAAGAAATTT 5340  
QY 5341 AATAGAAATAGTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 5400  
Db 5341 AATAGAAATAGTTTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 5400  
QY 5401 GTAATGTTCCAAATCTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5460  
Db 5401 GTAATGTTCCAAATCTAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5460  
QY 5461 AGCAATGTGAAAGTT 5475  
Db 5461 AGCAATGTGAAAGTT 5475

RESULT 2  
ABN89594  
ID ABN89594 standard; cDNA; 6525 BP.  
XX  
AC ABN89594;  
XX  
DT 18-SEP-2002 (first entry)  
XX  
DE Human ATP-binding cassette transporter ABCA5 cDNA SEQ ID NO:1.  
XX  
KW Human; ABCA5; ABCA6; ABCA9; ABCA10; ATP-binding cassette transporter;  
KW chromosome 17; chromosome 17q; chromosome 17q24; arteriosclerotic;  
KW gene therapy; cholesterol; lipophilic molecule; inflammation;  
KW prostaglandin; prostacyclin; arteriosclerosis; transport; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200246458-A2.  
XX  
PD 13-JUN-2002.  
XX  
PF 07-DEC-2001; 2001WO-EP015401.  
XX  
PR 07-DEC-2000; 2000EP-00403440.  
PR 23-JAN-2001; 2001US-0263231P.  
XX  
PA (AVET ) AVENTIS PHARMA SA.  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Derofle P, Rosier-Montus M, Prades C, Arnould-Reguigne I;  
PI Duverger N, Allikmets R, Dean M;  
XX  
DR WPI; 2002-557584/59.  
DR P-PSDB; ABB81574.  
XX  
PT A novel nucleic acid corresponding to ATP-binding cassette transporter  
PT genes and the encoded polypeptide, useful for preventing or treating a  
PT dysfunction in reverse transport of cholesterol.

XX PS Claim 1; Page 151-153; 216pp; English.

XX CC The present invention describes human ATP-binding cassette transporters

CC (ABC). Specifically described are the human ABCA5, ABCA6, ABCA9 and

CC ABCA10 genes (see ABN89594 to ABN89597) which encode the proteins given

CC in ABN83574 to ABN81577). ABN89598 to ABN89715 represent ABCA5, ABCA6,

CC ABCA9 and ABCA10 nucleotide fragments; and ABN89716 to ABN89806 represent

CC primers for ABCA5, ABCA6, ABCA9 and ABCA10 genes which are used in the

CC exemplification of the present invention. The ABC sequences have

CC antiarteriosclerotic activities and can be used in gene therapy. ABC

CC sequences can be used in the manufacture of a medicament intended for the

CC prevention and/or treatment of a subject affected by a dysfunction in the

CC reverse transport of cholesterol. The ABC proteins are involved in the

CC reverse transport of cholesterol, in membrane transport of lipophilic

CC molecules, in particular inflammation mediating substance such as

CC prostaglandins and prostacyclins, or in any pathology whose candidate

CC chromosomal region is situated on chromosome 17. They are also useful for

CC the manufacture of a medicament intended for prevention of

CC arteriosclerosis in various forms. The ABCA5, ABCA6, ABCA9 and ABCA10

CC genes are located to chromosome 17, more specifically to the 17q24 locus

XX SQ Sequence 6525 BP; 1973 A; 1108 C; 1258 G; 2185 T; 0 U; 1 Other;

Query Match 98.1%; Score 5373.2; DB 6; Length 6525;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 5375; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 97 AGGTTTATTCAGAAACATGTCACGTGCAATAGGAGGTAGGAGTTGGAGACAGCA 156

DB 994 AGGTTTATTCAGAAACATGTCACGTGCAATAGGAGGTAGGAGTTGGAGACAGCA 1053

QY 157 GAACACTTCTACTGAAGAATTACTTAATTAATGAGAACCAAAAGAGTAGTGTTCAGG 216

DB 1054 GAACACTTCTACTGAAGAATTACTTAATTAATGAGAACCAAAAGAGTAGTGTTCAGG 1113

QY 217 AAATCTTTTTCACATTTTTTTTTTTTATTTGGTTAATTAATTAATGAGATGATGCCAA 276

DB 1114 AAATCTTTTTCACATTTTTTTTTTTTATTTGGTTAATTAATTAATGAGATGATGCCAA 1173

QY 277 ATAAGAATATGAGAGTGCCTATATATAGTCACTCAATCCTATGACAGTTTACTCTTT 336

DB 1174 ATAAGAATATGAGAGTGCCTATATATAGTCACTCAATCCTATGACAGTTTACTCTTT 1233

QY 337 CTAATCTAATTTCTGGATATCTCCAGTGCACTAATATTACAAGCAGCATCATGCAAGAAG 396

DB 1234 CTAATCTAATTTCTGGATATCTCCAGTGCACTAATATTACAAGCAGCATCATGCAAGAAG 1293

QY 397 TGTCTACTGATCATCTACCTGATGTCATTAATTAATGAGATATACAAATGAAAGAA 456

DB 1294 TGTCTACTGATCATCTACCTGATGTCATTAATTAATGAGATATACAAATGAAAGAA 1353

QY 457 TGTAAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAAGACTCCA 516

DB 1354 TGTAAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAAGACTCCA 1413

QY 517 TGTCTATGAATCTGCTTTTTTCTGATATGATTCAGATATCTCTATTTATATGGATT 576

DB 1414 TGTCTATGAATCTGCTTTTTTCTGATATGATTCAGATATCTCTATTTATATGGATT 1473

QY 577 CAAGAGCTGGCTGTTCAAATCATGTAGGCTGCTCAGTACTGCTCCTCAGGTTTCACAG 636

DB 1474 CAAGAGCTGGCTGTTCAAATCATGTAGGCTGCTCAGTACTGCTCCTCAGGTTTCACAG 1533

QY 637 TTTTACAAGCATCCATGATGTCGCAATATACAGTTGAAGACCAATGTTTCTCTTTGGA 696

DB 1534 TTTTACAAGCATCCATGATGTCGCAATATACAGTTGAAGACCAATGTTTCTCTTTGGA 1593

QY 697 AGGAGCTGGAGTCAACTAAGCTGTTTATATGCGAGAACTCCTGTTGTGAAGATAGATA 756

DB 1594 AGGAGCTGGAGTCAACTAAGCTGTTTATATGCGAGAACTCCTGTTGTGAAGATAGATA 1653

QY 757 CCTTTCCCGGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTCACTTTTGGATACT 816

DB 1654 CCTTTCCCGGAGGAGTAATTTTAAATATACCTAGTTATAGCAATTTTCACTTTTGGATACT 1713

QY 817 TTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAA 876

DB 1714 TTTTGGCAATTCATATCGTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAA 1773

QY 877 TGGGACTTCATGATACATGCTTTTGGCTTTCCCTGGGTTCTTCATATACAAAGTTAATTT 936

DB 1774 TGGGACTTCATGATACATGCTTTTGGCTTTCCCTGGGTTCTTCATATACAAAGTTAATTT 1833

QY 937 TTCTTATGTCCTCTCTTATGCGAGTCAATGCGACAGCTTCTTTGTTATTTCCCTCAAAGTA 996

DB 1834 TTCTTATGTCCTCTCTTATGCGAGTCAATGCGACAGCTTCTTTGTTATTTCTCTCAAAGTA 1893

QY 997 GCAGCATTTGTGATATTTCTGCTTTTTTCTTTATGAGTATCATCTGATTTTTCGCTT 1056

DB 1894 GCAGCATTTGTGATATTTCTGCTTTTTTCTTTATGAGTATCATCTGATTTTTCGCTT 1953

QY 1057 TAAATGCTGACACCTCTTTTAAATAATCAAAACATGCGGAATAGTTGAATTTTGTGTA 1116

DB 1954 TAAATGCTGACACCTCTTTTAAATAATCAAAACATGCGGAATAGTTGAATTTTGTGTA 2013

QY 1117 CTGTGGCTTTTGGATTTATTTGGCTTTATGATAATCCTCATAGAAAGTTTCCCAATCGT 1176

DB 2014 CTGTGGCTTTTGGATTTATTTGGCTTTATGATAATCCTCATAGAAAGTTTCCCAATCGT 2073

QY 1177 TAGTGTGGCTTTTCACTGCTTTCTGTCACGTACTTTTGTGATTGGTATTGCACAGTCA 1236

DB 2074 TAGTGTGGCTTTTCACTGCTTTCTGTCACGTACTTTTGTGATTGGTATTGCACAGTCA 2133

QY 1237 TGCATTTTAGAAGATTTTAAATGAAGGTGCTTCATTTTCAAAATTTGACTCGAGGGCCATATC 1296

DB 2134 TGCATTTTAGAAGATTTTAAATGAAGGTGCTTCATTTTCAAAATTTGACTCGAGGGCCATATC 2193

QY 1297 CTCTAATTTATACAAATATCATGCTGCACCTTAATAGTATTTCTATGCTCTCTGGCTG 1356

DB 2194 CTCTAATTTATACAAATATCATGCTGCACCTTAATAGTATTTCTATGCTCTCTGGCTG 2253

QY 1357 TCTATCTTGAATCAAGTCAATTTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 1416

DB 2254 TCTATCTTGAATCAAGTCAATTTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 2313

QY 1417 TGAAGCTTTCATATTTGGTCAAAGAGCAAAAGAAATTAAGAGGAGTTATCAGAGGGCAATG 1476

DB 2314 TGAAGCTTTCATATTTGGTCAAAGAGCAAAAGAAATTAAGAGGAGTTATCAGAGGGCAATG 2373

QY 1477 TTAATGAAATATTTAGTTTTTGTAGTGAATTTTTCAGCCAGTTCTTCTCAGAAATTTGTAGGAA 1536

DB 2374 TTAATGAAATATTTAGTTTTTGTAGTGAATTTTTCAGCCAGTTCTTCTCAGAAATTTGTAGGAA 2433

QY 1537 AAGAAGCCATPAAGAAATTTAGTGGTATTTCAGAAAGACATACAGAAAGAGGGGTGAATAATGTGG 1596

DB 2434 AAGAAGCCATPAAGAAATTTAGTGGTATTTCAGAAAGACATACAGAAAGAGGGGTGAATAATGTGG 2493

QY 1597 AGGCTTTGAGAAATTTGTCAATTTGACATATATCAGAGGTCAAGTACTGCTTACTTTGGCC 1656

DB 2494 AGGCTTTGAGAAATTTGTCAATTTGACATATATCAGAGGTCAAGTACTGCTTACTTTGGCC 2553

QY 1657 ACAGTGGAAACAGGAAAGAGTACATTTGATGAATATCTTTGTGGACTCTGCCACCTTCTG 1716

DB 2554 ACAGTGGAAACAGGAAAGAGTACATTTGATGAATATCTTTGTGGACTCTGCCACCTTCTG 2613

QY 1717 ATGGGTTTGCATCTATATATATGACACAGGTCTCAGAAATAGATGAATGTTTGAAGCAA 1776

DB 2614 ATGGGTTTGCATCTATATATATGACACAGGTCTCAGAAATAGATGAATGTTTGAAGCAA 2673

QY 1777 GAAAAATGATTTGGCAATTTGTCCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAAG 1836

DB 2674 GAAAAATGATTTGGCAATTTGTCCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAAG 2733

QY 1837 AAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCCAACATATATATCAAGAG 1896

Db 2734 AAAATTATCAATTTTGGGTTCAATCAAAAGGGATACCAGCCAAACATATAATACAAAGAG 2793  
Qy 1897 TGCAAGAGGTTTACTAGATTTAGACATGCAGACTATCAAGATACCAAGCTAAAAAAT 1956  
Db 2794 TGCAGAGGTTTACTAGATTTAGACATGCAGACTATCAAGATACCAAGCTAAAAAAT 2853  
Qy 1957 TAAGTGGTGTCAAAAAGAAAGCTGTCAATTAGGAATTCCTGTTCTTGGGAACCCAAAGA 2016  
Db 2854 TAAGTGGTGTCAAAAAGAAAGCTGTCAATTAGGAATTCCTGTTCTTGGGAACCCAAAGA 2913  
Qy 2017 TACTGCTGTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGA 2076  
Db 2914 TACTGCTGTAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGA 2973  
Qy 2077 ATCTTTTAAAATACAGAAAGCCCAATCCGGGTGACAGTGTTCAGTACTCATTTTCATGGAATG 2136  
Db 2974 ATCTTTTAAAATACAGAAAGCCCAATCCGGGTGACAGTGTTCAGTACTCATTTTCATGGAATG 3033  
Qy 2137 AAGCTGACATCTTTGAGATAGGAAGCTGTGATATCAAGAAATGCTGAAATGTTG 2196  
Db 3034 AAGCTGACATCTTTGAGATAGGAAGCTGTGATATCAAGAAATGCTGAAATGTTG 3093  
Qy 2197 GTTCTTCAATGTTCTTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
Db 3094 GTTCTTCAATGTTCTTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3153  
Qy 2257 ACAAATATTTGGCCACAGATCTCTTCTTCACTGGTTAAACAAACATATACCTGGAGCTA 2316  
Db 3154 ACAAATATTTGGCCACAGATCTCTTCTTCACTGGTTAAACAAACATATACCTGGAGCTA 3213  
Qy 2317 CTTTATTACAAAGAAATGACCAACAACTTTGTGTATAGCTTGGCTTTCAAGGACATGGACA 2376  
Db 3214 CTTTATTACAAAGAAATGACCAACAACTTTGTGTATAGCTTGGCTTTCAAGGACATGGACA 3273  
Qy 2377 AATTTTCAGGTTTGTCTTCTGGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATG 2436  
Db 3274 AATTTTCAGGTTTGTCTTCTGGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATG 3333  
Qy 2437 GTGTTTCCATGACGACTTTTGGAGAGCGTATTTTAAAGCTAGAAGTTGAAGCAGAAATG 2496  
Db 3334 GTGTTTCCATGACGACTTTTGGAGAGCGTATTTTAAAGCTAGAAGTTGAAGCAGAAATG 3393  
Qy 2497 ACCAAGAGATATAGTATTTTACTCAGAGCCACTGGAGGAAGAAATGGATTCAAAAT 2556  
Db 3394 ACCAAGAGATATAGTATTTTACTCAGAGCCACTGGAGGAAGAAATGGATTCAAAAT 3453  
Qy 2557 CTTTTCATGAATGGAACAGAGCTTACTTATTTCTTCAAAACCAAGGCTTCTAGTGA 2616  
Db 3454 CTTTTCATGAATGGAACAGAGCTTACTTATTTCTTCTGAACCAAGGCTTCTAGTGA 3513  
Qy 2617 GCACCATGAGCCTTTGGAAAACAAACAGATGTATACAAATAGCAAGTTTCAATTTCTTTACCT 2676  
Db 3514 GCACCATGAGCCTTTGGAAAACAAACAGATGTATACAAATAGCAAGTTTCAATTTCTTTACCT 3573  
Qy 2677 TGAACGTAAGCTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 2736  
Db 3574 TGAACGTAAGCTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAG 3633  
Qy 2737 TTCAGATTTTATGTTTTTGGTTTCATCACTCTTTTAAAAATGCTGTGGTTCCCATCAAC 2796  
Db 3634 TTCAGATTTTATGTTTTTGGTTTCATCACTCTTTTAAAAATGCTGTGGTTCCCATCAAC 3693  
Qy 2797 TTGTTCCAGACTTATTTTCTTAAACCTGGAGCAAAACCAATATAACAAACAAAGTC 2856  
Db 3694 TTGTTCCAGACTTATTTTCTTAAACCTGGAGCAAAACCAATATAACAAACAAAGTC 3753  
Qy 2857 TGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAAGCC 2916  
Db 3754 TGCTTCTTCAAAATCTGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAAGCC 3813  
Qy 2917 AGAACATTAATGGTGAGATGAATTAATGACAGTCACTATGTATCCGTCCTCCCATAGT 2976  
Db 3814 AGAACATTAATGGTGAGATGAATTAATGACAGTCACTATGTATCCGTCCTCCCATAGT 3873

Qy 2977 CGGCTTTAAATGTGATGCATTCAGAAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTA 3036  
Db 3874 CGGCTTTAAATGTGATGCATTCAGAAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTA 3933  
Qy 3037 CTATGGTTTAAATCTTTTACCTATATATTAGTGAATCATTAAGTAACTACTACTATCTTTATCAT 3096  
Db 3934 CTATGGTTTAAATCTTTTACCTATATATTAGTGAATCATTAAGTAACTACTACTATCTTTATCAT 3993  
Qy 3097 TAAATGTGACTGAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTAAGTGA 3156  
Db 3994 TAAATGTGACTGAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTAAGTGA 4053  
Qy 3157 TAGTTTTTAAATTTAGCTGTATTTTCAAGCAGCTTTTGGTGAATCATTTGTTACTGCAA 3216  
Db 4054 TAGTTTTTAAATTTAGCTGTATTTTCAAGCAGCTTTTGGTGAATCATTTGTTACTGCAA 4113  
Qy 3217 TGCACACCTTACTTTGCCATGGAATAATGCAGAGAAATCATTAAGTCAAAAGCTTATTAAC 3276  
Db 4114 TGCACACCTTACTTTGCCATGGAATAATGCAGAGAAATCATTAAGTCAAAAGCTTATTAAC 4173  
Qy 3277 TTAACCTTTCAGGTCTTTTGGCAATCTGCATATTGGATTGGCAAGCTGTTGTTGATATCC 3336  
Db 4174 TTAACCTTTCAGGTCTTTTGGCAATCTGCATATTGGATTGGCAAGCTGTTGTTGATATCC 4233  
Qy 3337 CCTTATTTTTTATCATTTCTTATTTTGCATGCTAGGAAGCTTATTTGGCATTTTCAATTATGGAT 3396  
Db 4234 CCTTATTTTTTATCATTTCTTATTTTGCATGCTAGGAAGCTTATTTGGCATTTTCAATTATGGAT 4293  
Qy 3397 TATATTTTTTATCATTTAAAGTTCCTTGTGCTGTGGTTTTTGGCTTATTTGGTTATGTTCCAT 3456  
Db 4294 TATATTTTTTATCATTTAAAGTTCCTTGTGCTGTGGTTTTTGGCTTATTTGGTTATGTTCCAT 4353  
Qy 3457 CAGTTATTTCTGTTCACTTATATTGCTTCTTCACTTTAAGAAAATTTTAAATACCAAAAG 3516  
Db 4354 CAGTTATTTCTGTTCACTTATATTGCTTCTTCACTTTAAGAAAATTTTAAATACCAAAAG 4413  
Qy 3517 AATTTTGGCTCATTTATCTATTCTGTGGCAGCGTGTGATTTGCAATCACTGAAATAA 3576  
Db 4414 AATTTTGGCTCATTTATCTATTCTGTGGCAGCGTGTGATTTGCAATCACTGAAATAA 4473  
Qy 3577 CTTTCTTTTATGGGATACACAAATTCGAATTTCTTCAATATGCTTTTGTATCATCATTC 3636  
Db 4474 CTTTCTTTTATGGGATACACAAATTCGAATTTCTTCAATATGCTTTTGTATCATCATTC 4533  
Qy 3637 CAATCTATCCACTTCTAGGTTGCTGATTTCTTTCAATAAGATTTCTTGGAGAATGTAC 3696  
Db 4534 CAATCTATCCACTTCTAGGTTGCTGATTTCTTTCAATAAGATTTCTTGGAGAATGTAC 4593  
Qy 3697 GAAAAAATGGACACCTATAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATCGCCTT 3756  
Db 4594 GAAAAAATGGACACCTATAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATCGCCTT 4653  
Qy 3757 ACCTGCACTGTACTGTGGATTTTCTCTTACAATACTATGAGAAAAAATATGGAGGCA 3816  
Db 4654 ACCTGCACTGTACTGTGGATTTTCTCTTACAATACTATGAGAAAAAATATGGAGGCA 4713  
Qy 3817 GATCAATTAAGAAAGATCCCTTTTCAGAAACCTTTTCAACGAGTCTTAAATAATAGGAAGC 3876  
Db 4714 GATCAATTAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAGTCTTAAATAATAGGAAGC 4773  
Qy 3877 TTCCAGAACCAACAGACAAATGAGGATGAAGATGTCAAAAGCTGAAAGACTAAAGG 3936  
Db 4774 TTCCAGAACCAACAGACAAATGAGGATGAAGATGTCAAAAGCTGAAAGACTAAAGG 4833  
Qy 3937 TCAAAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAACCATTCATTTATGGTCAGCAAT 3996  
Db 4834 TCAAAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAACCATTCATTTATGGTCAGCAAT 4893  
Qy 3997 TGCATAAAGAAATATGATGACAAAGAAATTTTCTTCTTCAAGAAAAGTAAAGAAAGTGG 4056  
Db 4894 TGCATAAAGAAATATGATGACAAAGAAATTTTCTTCTTCAAGAAAAGTAAAGAAAGTGG 4953

QY 4057 CAACTAAATACATCTCTTCTGTGTGAAAAAGAGAGATCTTAGAGACTATTGGGTCCAA 4116  
DB 4954 CAACTAAATACATCTCTTCTGTGTGAAAAAGAGAGATCTTAGAGACTATTGGGTCCAA 5013  
QY 4117 ATGGTGTCTGGCAAAAGCAAAATTAATTAATCTGGTGGTGATATTGAACCACTTCAG 4176  
DB 5014 ATGGTGTCTGGCAAAAGCAAAATTAATTAATCTGGTGGTGATATTGAACCACTTCAG 5073  
QY 4177 GCCAGTATTTTAGAGAGATTAATCTTCAGAGACAAGTGAAGATGATGATCACTGAAGT 4236  
DB 5074 GCCAGTATTTTAGAGAGATTAATCTTCAGAGACAAGTGAAGATGATGATCACTGAAGT 5133  
QY 4237 GTATGGTGTCTGGTCTCAGATAAACCCTTGTGGCCAGATACATCTTCAGAGAACATT 4296  
DB 5134 GTATGGTGTCTGGTCTCAGATAAACCCTTGTGGCCAGATACATCTTCAGAGAACATT 5193  
QY 4297 TTGAATTTATGGAGCTGTCAAGGAATGAGTGAAGTGAACATGAAGAAAGTCAATAGTC 4356  
DB 5194 TTGAATTTATGGAGCTGTCAAGGAATGAGTGAAGTGAACATGAAGAAAGTCAATAGTC 5253  
QY 4357 GAATACACATCATCTTGAATTTAAAGACATCTTCAGAGACTGTAAAGAACTACCTG 4416  
DB 5254 GAATACACATCATCTTGAATTTAAAGACATCTTCAGAGACTGTAAAGAACTACCTG 5313  
QY 4417 CAGGAATCAAGCAAGTGTGTCTTAAGTATGCTAGGGAATCCTCAGATTAATT 4476  
DB 5314 CAGGAATCAAGCAAGTGTGTCTTAAGTATGCTAGGGAATCCTCAGATTAATT 5373  
QY 4477 TCGTAGATGAACATCTCAGATGATGGATCCCAAGCCCAACAGCACATGTGGCGAGCAA 4536  
DB 5374 TCGTAGATGAACATCTCAGATGATGGATCCCAAGCCCAACAGCACATGTGGCGAGCAA 5433  
QY 4537 TTGCAACTCAATTTAAAAACAGAAAGCGGCTGTATCTGACCACTCACTATATGAGG 4596  
DB 5434 TTGCAACTCAATTTAAAAACAGAAAGCGGCTGTATCTGACCACTCACTATATGAGG 5493  
QY 4597 AGCAGAGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGAGTGAAGTATATCG 4656  
DB 5494 AGCAGAGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGAGTGAAGTATATCG 5553  
QY 4657 GAACAGTACCAATCTAAAGAGTAAATTTGAAAAAGGCTACTTTTGGAAATTAATGA 4716  
DB 5554 GAACAGTACCAATCTAAAGAGTAAATTTGAAAAAGGCTACTTTTGGAAATTAATGA 5613  
QY 4717 AGGACTGGATAGAAAACCTPAGAAGTAGACCGCTTCAAAGGAAATTCAGTATATTTCC 4776  
DB 5614 AGGACTGGATAGAAAACCTPAGAAGTAGACCGCTTCAAAGGAAATTCAGTATATTTCC 5673  
QY 4777 CAATGCAAGCGCTCAGGAAGCTTTTCTCTATTTTGGCTTATAAATTCCTAAGGAAG 4836  
DB 5674 CAATGCAAGCGCTCAGGAAGCTTTTCTCTATTTTGGCTTATAAATTCCTAAGGAAG 5733  
QY 4837 ATGTTCAAGTCCCTTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCCA 4896  
DB 5734 ATGTTCAAGTCCCTTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCCA 5793  
QY 4897 TTGAAGATATAGCTTTTCTCAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAG 4956  
DB 5794 TTGAAGATATAGCTTTTCTCAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAG 5853  
QY 4957 AACAGAGGAGGAAGATAATAGTTGTGGAACCTTTAAACAGCACATTTTGGTGGGAACGAA 5016  
DB 5854 AACAGAGGAGGAAGATAATAGTTGTGGAACCTTTAAACAGCACATTTTGGTGGGAACGAA 5913  
QY 5017 CACAGAAATAGAGTAGTATTTGAAATTTGTTGTTGCTGCTCTACTGGGACTTCT 5076  
DB 5914 CACAGAAATAGAGTAGTATTTGAAATTTGTTGTTGCTGCTCTACTGGGACTTCT 5973  
QY 5077 TTCTTTTTCACATTAATTTAACTTTGTTTAAAGTTTTTATTGGAATGTTAACTGGA 5136  
DB 5974 TTCTTTTTCACATTAATTTAACTTTGTTTAAAGTTTTTATTGGAATGTTAACTGGA 6033  
QY 5137 GRACCAAGAACCGCACTTGAATTTTCTTAAGCTCTTAATTAAGATGCTGTGTGTG 5196

DB 6034 GAACCAAGAACGCACTTGAAATTTTCTAAGCTCCCTAATTAATGAATGCTGTGTGTG 6093  
QY 5197 TTTTGTCTTTCTTTAAATAAAGCTATGATATAATTAAGTGAAGCTGCATGTTTGTATTGA 5256  
DB 6094 TTTTGTCTTTCTTTAAATAAAGCTATGATATAATTAAGTGAAGCTGCATGTTTGTATTGA 6153  
QY 5257 AGTATATTGAACATATAGTTTGTATGATCATCTTTTCCACCTCAGAACAGTGTCTCT 5316  
DB 6154 AGTATATTGAACATATAGTTTGTATGATCATCTTTTCCACCTCAGAACAGTGTCTCT 6213  
QY 5317 GAATTTGTGATTTTAAAGGAATTTGAATAGAATAGTTTATTTTAAAGTTATCTTTAAGTT 5376  
DB 6214 GAATTTGTGATTTTAAAGGAATTTGAATAGAATAGTTTATTTTAAAGTTATCTTTAAGTT 6273  
QY 5377 TATGCCATCTTCTTAATAAGTACGTAATGTTCCATCTTAATAAATAAATACTAATTA 5436  
DB 6274 TATGCCATCTTCTTAATAAGTACGTAATGTTCCATCTTAATAAATAAATACTAATTA 6333  
QY 5437 CTATGTCATAGAAAGATACATAAAGCAATGTGAAAGTT 5475  
DB 6334 CTATGTCATAGAAAGATACATAAAGCAATGTGAAAGTT 6372

RESULT 3  
ABS57749  
ID ABS57749 standard; cDNA; 5463 BP.  
XX ABS57749;  
AC ABS57749;  
XX AC  
XX AC  
DT 04-FEB-2003 (first entry)  
XX  
DE cDNA encoding novel human ATP binding cassette ABCA5 transporter #1.  
XX Human; ATP binding cassette; ABC; ABCA5; transporter;  
XX Neurotoxin transport; beta-amyloid peptide; chromosome mapping;  
XX blood brain barrier transport; tissue typing; predictive medicine;  
XX ABCA5 mediated disorder; ABCA5 related disorder; gene therapy; gene; ss.  
OS Homo sapiens.  
Key Location/Qualifiers  
FH 114..5030  
FT /tag= a  
FT /product= "ABCA5"  
FT /note= "ATP binding cassette (ABC) A5 transporter"  
FT /transl\_except= (pos:3273..3275, aa:Ser)  
XX  
PN US2002123107-A1.  
XX  
XX 05-SEP-2002.  
XX 01-MAR-2002; 2002US-00090458.  
XX  
XX 02-MAR-2001; 2001US-0272885P.  
XX  
XX (ACT1-) ACTIVE PASS PHARM INC.  
XX  
XX Chen H, Kilinski L, Le Bihan S;  
XX WPI; 2003-066798/06.  
XX P-PSDB; ABG72423.  
XX  
XX Novel isolated ATP binding cassette transporter family polypeptide,  
XX ABCA5, useful for treating disorders associated with aberrant or unwanted  
XX ABCA5 transporter expression or activity.  
XX  
XX Claim 2; Fig 1; 52pp; English.  
XX  
XX The invention describes an isolated ATP binding cassette (ABC)  
XX transporter family polypeptide (I), designated ABCA5. (i) or the  
XX polynucleotide encoding it (II) are useful as targets for developing  
XX modulating agents to regulate a variety of cellular processes,

CC particularly the transport of neurotoxic molecules, e.g., beta-amyloid  
 CC peptide (Abeta), across cell membranes or, e.g., the blood brain barrier  
 CC (BBB), as targets for developing modulating agents of multi-drug  
 CC resistance, as diagnostic and therapeutic tools, or to treat disorders  
 CC associated with aberrant or unwanted ABCA5 transporter expression or  
 CC activity. (I), (II) or a host cell (III) expressing (II) are useful in  
 CC screening assays, detection assays (e.g., chromosomal mapping, tissue  
 CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,  
 CC prognostic assays, monitoring clinical trials and pharmacogenomics), and  
 CC in methods of treatment (e.g., therapeutic and prophylactic). (I) or  
 CC (III) are useful as reagents or targets in assays applicable to treatment  
 CC and diagnosis of ABCA5-mediated or related disorders (I). Is useful to  
 CC screen for naturally occurring ABCA5 substrates; to screen for drugs or  
 CC compounds which modulate ABCA5 activity; as a bait protein in a yeast two  
 CC -hybrid or three-hybrid assay; and to identify other proteins which bind  
 CC to or interact with ABCA5. (II) Is useful in: gene therapy; to detect  
 CC ABCA5 mRNA or a genetic alteration in a ABCA5 gene; to modulate ABCA5  
 CC activity; to locate gene regions associated with genetic disease or to  
 CC associate ABCA5 with the disease, to identify an individual from a minute  
 CC biological sample (tissue typing), and to aid in forensic identification  
 CC of the biological sample. this sequence encodes a novel human APO binding  
 CC cassette (ABC) A5 transporter  
 XX

SQ Sequence 5463 BP; 1699 A; 912 C; 1055 G; 1794 T; 0 U; 3 Other;

Query Match 97.8%; Score 5356.6; DB 7; Length 5463;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 5440; Conservative 0; Mismatches 4; Indels 50; Gaps 3;  
 1 GGGTCGGCCCTCGCACAGATCCNAGCTGGTGCACCGCAGCTGAGTCAACAGACTCGAGC 60  
 1 GGGTCGGCCCTCGCACAGATCCNAGCTGGTGCACCGCAGCTGAGTCAACAGACTCGAGC 60  
 61 GGGTCGGCCCTCGCACAGCTCTCGGCTCGGCGCTCGGCGCTGAGTTCAGAAACATGTCCTA 120  
 61 GGGTCGGCCCTCGCACAGCTCTCGGCTCGGCGCTCGGCGCTGAGTTCAGAAACATGTCCTA 120  
 121 CTGCAATTAGGAGGTAGGAGTTCGGAGACAGACAGACACTTCTACTGAAGAATTACT 180  
 121 CTGCAATTAGGAGGTAGGAGTTCGGAGACAGACAGACACTTCTACTGAAGAATTACT 180  
 181 TAATTAATGCGAGACCAAAAGAGTAGTGTTCAGGAATTCCTTTCCACTATTTTTT 240  
 181 TAATTAATGCGAGACCAAAAGAGTAGTGTTCAGGAATTCCTTTCCACTATTTTTT 240  
 241 TATTTGGTAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 241 TATTTGGTAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300  
 301 ATATAGAACTCAATCTATGGAAGAGTTCCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360  
 301 ATATAGAACTCAATCTATGGAAGAGTTCCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360  
 361 CAGTGACTAATTAATCAAGCAGATCATGACAGAGTGTCTACTGATCATCTACTGATG 420  
 361 CAGTGACTAATTAATCAAGCAGATCATGACAGAGTGTCTACTGATCATCTACTGATG 420  
 421 TCATTAATTAATGAGAAATATCAAAATGAAAGAAATGTTAACTCCAGTCTCTTAAGC 480  
 421 TCATTAATTAATGAGAAATATCAAAATGAAAGAAATGTTAACTCCAGTCTCTTAAGC 480  
 481 CGAGCAACTTCTAGGTGGTTCAGAGCTCCATGCTCTATGAACTTCGTTTTTTTC 540  
 481 CGAGCAACTTCTAGGTGGTTCAGAGCTCCATGCTCTATGAACTTCGTTTTTTTC 540  
 541 CTGATATGATTCAGGATCTTCTTATTTATGATGATTCAGAGCTGGCTGTTCAAAATCAT 600  
 541 CTGATATGATTCAGGATCTTCTTATTTATGATGATTCAGAGCTGGCTGTTCAAAATCAT 600  
 601 GTGAGGCTGCTAGTACTGCTCTAGGTTTCAAGTTTACAGATCCATAGATGCTG 660  
 601 GTGAGGCTGCTAGTACTGCTCTAGGTTTCAAGTTTACAGATCCATAGATGCTG 660

QY 661 CCATTATACAGTTGAAGACCAATGTTTCTCTTTGGAAGAGCTGGAGTCAACTAAAGCTG 720  
 DB 661 CCATTATACAGTTGAAGACCAATGTTTCTCTTTGGAAGAGCTGGAGTCAACTAAAGCTG 720  
 QY 721 TTATTATGGAGAACTGCTGTGTAGTAATAGATACCTTTCCCGAGAGTAATTTAA 780  
 DB 721 TTATTATGGAGAACTGCTGTGTAGTAATAGATACCTTTCCCGAGAGTAATTTAA 780  
 QY 781 TATACCTAGTTATAGCAATTTTCACTTTTGGATCTTTTGGCAATTCATATGTCAGAG 840  
 DB 781 TATACCTAGTTATAGCAATTTTCACTTTTGGATCTTTTGGCAATTCATATGTCAGAG 840  
 QY 841 AAAAAAGAAAAAATAAAGAAATTTTAAAGTAATGGACTTCATGATCTGCTTTT 900  
 DB 841 AAAAAAGAAAAAATAAAGAAATTTTAAAGTAATGGACTTCATGATCTGCTTTT 900  
 QY 901 GGCTTTCCCTGGGTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCTTTTATGGCAG 960  
 DB 901 GGCTTTCCCTGGGTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCTTTTATGGCAG 960  
 QY 961 TCATTGCGACAGCTTCTTTGTTTATTTCTTCAAGTAGCAGCATTTGATATTTCTGCTTT 1020  
 DB 961 TCATTGCGACAGCTTCTTTGTTTATTTCTTCAAGTAGCAGCATTTGATATTTCTGCTTT 1020  
 QY 1021 TTTTCTTTATGATTTATCTGTTATTTTTCCTTTTAAATGCTGACACCTCTTTTAAAAA 1080  
 DB 1021 TTTTCTTTATGATTTATCTGTTATTTTTCCTTTTAAATGCTGACACCTCTTTTAAAA 1080  
 QY 1081 AATCAAAAATGTTGGGAATGTTGAAATTTTGTCTGCTGCTTTGGAATTTATGGCC 1140  
 DB 1081 AATCAAAAATGTTGGGAATGTTGAAATTTTGTCTGCTGCTTTGGAATTTATGGCC 1140  
 QY 1141 TTAATGATTAATCTCTATAGAAAGTTTCCCAATCGTTAGTGGCTTTTTCAGTCTTTCT 1200  
 DB 1141 TTAATGATTAATCTCTATAGAAAGTTTCCCAATCGTTAGTGGCTTTTTCAGTCTTTCT 1200  
 QY 1201 GTCACTGTATTTTGTGATTTGATTTGACAGCTCATGATTTAGAGATTTTATGAAG 1260  
 DB 1201 GTCACTGTATTTTGTGATTTGATTTGACAGCTCATGATTTAGAGATTTTATGAAG 1260  
 QY 1261 GTGCTTTCAATTTTCAATTTGACAGGCCCCATATCTCTTAATTTTCAATTTATCATGC 1320  
 DB 1261 GTGCTTTCAATTTTCAATTTGACAGGCCCCATATCTCTTAATTTTCAATTTATCATGC 1320  
 QY 1321 TCACACTTAATAGTATATTTCTATGCTCTGCTGCTCTATCTGATCAAGTCAATTCAG 1380  
 DB 1321 TCACACTTAATAGTATATTTCTATGCTCTGCTGCTCTATCTGATCAAGTCAATTCAG 1380  
 QY 1381 GGGAAATTTGGCTTACCGAGATCATCTTTATATTTTCTGAGGCTTTCATATTTGGTCAA 1440  
 DB 1381 GGGAAATTTGGCTTACCGAGATCATCTTTATATTTTCTGAGGCTTTCATATTTGGTCAA 1440  
 QY 1441 GCAAAAGAAATTTATGAGGAGTTATCAGAGGCAATGTTAATGGAATATTAGTTTATG 1500  
 DB 1441 GCAAAAGAAATTTATGAGGAGTTATCAGAGGCAATGTTAATGGAATATTAGTTTATG 1500  
 QY 1501 AAATTTATGAGCCAGTTCTTCTAGATTTGTTAGAAAGAGCCATTAAGATTTAGTGGTA 1560  
 DB 1501 AAATTTATGAGCCAGTTCTTCTAGATTTGTTAGAAAGAGCCATTAAGATTTAGTGGTA 1560  
 QY 1561 TCCAGAGACATACAGAAAGAGGGTGAATAATGTTGAGGCTTTTGAAATTTGTCTATTG 1620  
 DB 1561 TCCAGAGACATACAGAAAGAGGGTGAATAATGTTGAGGCTTTTGAAATTTGTCTATTG 1620  
 QY 1621 ACATATATGAGGTCAGATTTACTGCTTACTTGGCCACAGTGGAAAGAGAGTACAT 1680  
 DB 1621 ACATATATGAGGTCAGATTTACTGCTTACTTGGCCACAGTGGAAAGAGAGTACAT 1680  
 QY 1681 TGAATGAATATTTCTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATATATGGAC 1740  
 DB 1681 TGAATGAATATTTCTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATATATGGAC 1740  
 QY 1741 ACAGAGTCTCAGAAATAGATGAATAATGTTTGAAGAGAGAAATATGATTTGCTCCAC 1800

1741 ACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATGGCATTTGTCCAC 1800  
1801 AGTTAGATATACACTTTTCATGTTTTCAGTAGAGAGAAAAATTTATCAATTTTGCTTCAA 1860  
1801 AGTTAGATATACACTTTTCATGTTTTCAGTAGAGAGAAAAATTTATCAATTTTGCTTCAA 1860  
1861 TCAAAGGATACCAAGCCAAACATATATACAAAGAGTGCAGAGGTTTATCTAGATTAG 1920  
1921 ACATGCGACATATCAAGATACCAAGCTTAAATTAAGTGGTGGTCAAAAAAGAAC 1980  
1921 ACATGCGACATATCAAGATACCAAGCTTAAATTAAGTGGTGGTCAAAAAAGAAC 1980  
1981 TGTCATTAGGAATTCCTGTTCTGGGAACCCAAAGATACCTGCTAGATGAACCAACAG 2040  
1981 TGTCATTAGGAATTCCTGTTCTGGGAACCCAAAGATACCTGCTAGATGAACCAACAG 2040  
2041 CTGGAATGGAACCCCTGTTCTGACATATGTATGGAACTTTTAAATACAGAAAGCCA 2100  
2041 CTGGAATGGAACCCCTGTTCTGACATATGTATGGAACTTTTAAATACAGAAAGCCA 2100  
2101 ATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTGACATAGGA 2160  
2101 ATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTGACATAGGA 2160  
2161 AAGCTGTGATATCAAGAAATGCTGAAATGTGTTGGTTCCTCAATGTTCTCAAAAGTA 2220  
2161 AAGCTGTGATATCAAGAAATGCTGAAATGTGTTGGTTCCTCAATGTTCTCAAAAGTA 2220  
2221 AATGGGGATCGGCTACCGCTGACATGTATACATAGACAAATATTTGTCACAGAACTTC 2280  
2221 AATGGGGATCGGCTACCGCTGACATGTATACATAGACAAATATTTGTCACAGAACTTC 2280  
2281 TTTCTTCACTGTTTAAACAAACATATACCTGGAGCTACTTTATTAACACAGAAATGACCAAC 2340  
2281 TTTCTTCACTGTTTAAACAAACATATACCTGGAGCTACTTTATTAACACAGAAATGACCAAC 2340  
2341 AACTTGTGTATAGTGTGCTTCAAGGACATGGAACAAATTTTCAGGTTGTTTCTGCCC 2400  
2341 AACTTGTGTATAGTGTGCTTCAAGGACATGGAACAAATTTTCAGGTTGTTTCTGCCC 2400  
2401 TAGACAGTCAATCAAAATTTGGTGTCTATTTCTATGTTGTTTCCATGACGACTTTGGAAG 2460  
2401 TAGACAGTCAATCAAAATTTGGTGTCTATTTCTATGTTGTTTCCATGACGACTTTGGAAG 2460  
2461 AGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACACAGCAGATTTATAGTGTATTTA 2520  
2461 AGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACACAGCAGATTTATAGTGTATTTA 2520  
2521 CTCAGAGCCACTGGAGGAAGAAATGGATTTCAAAATCTTTTATGAAATGGAACAGAGCT 2580  
2521 CTCAGAGCCACTGGAGGAAGAAATGGATTTCAAAATCTTTTATGAAATGGAACAGAGCT 2580  
2581 TACTTATTTCTTCTGAAACCAAGGCTTCTTAGTGAGCACCATGAGCCCTTTGGAAACAAAC 2640  
2581 TACTTATTTCTTCTGAAACCAAGGCTTCTTAGTGAGCACCATGAGCCCTTTGGAAACAAAC 2640  
2641 AAGTGTATCAATAGCAAAAGTTTCAATTTCTTTACCTTGAAACGTTAAATCAGTGA 2700  
2641 AAGTGTATCAATAGCAAAAGTTTCAATTTCTTTACCTTGAAACGTTAAATCAGTGA 2700  
2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAAGTTTATGTTTGGTTC 2760  
2701 GATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAAGTTTATGTTTGGTTC 2760  
2761 ATCACTCTTTTAAATGCTGTGTTTCCATCAAACTTTGTCACAGCTTATATTTCTTAA 2820  
2761 ATCACTCTTTTAAATGCTGTGTTTCCATCAAACTTTGTCACAGCTTATATTTCTTAA 2820  
2821 AACCTGGAGACAAACCAATAAATACAAACAAAGTCTGCTTCTTCAAAATTTCTGCT---- 2876

2821 AACCTGGAGACAAACCAATAAATACAAACAAAGTCTGCTTCTTCAAAATTTCTGCTGGTG 2880  
2877 -----GACTCAGATATACAGTATCTTATAGCTTTTTCACAAAGCCAGAAACA 2922  
2881 AGAGTGTNNGTGAAGACTCAGATATACAGTATCTTATAGCTTTTTCACAAAGCCAGAAACA 2940  
2923 TAATGCTGACGATGATTAAATGACAGTACTATGTAATCGTGGCTCCCATAGTGGGCTT 2982  
2941 TAATGCTGACGATGATTAAATGACAGTACTATGTAATCGTGGCTCCCATAGTGGGCTT 3000  
2983 TAAATGTGATGCAATTCAGAAAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTACTATGG 3042  
3001 TAAATGTGATGCAATTCAGAAAAAGGACTATGTTTTTTCAGCTGTTTTTCAACAGTACTATGG 3060  
3043 TTTATTTCTTTACTATATTTAGTGAATATCATTTAGTAACTACTATCTTTTATCATTTAAATG 3102  
3061 TTTATTTCTTTACTATATTTAGTGAATATCATTTAGTAACTACTATCTTTTATCATTTAAATG 3120  
3103 TGACTGAAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAGAAATTTACTGATATAGTTT 3162  
3121 TGACTGAAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAGAAATTTACTGATATAGTTT 3180  
3163 TTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGG- AATCATGTTTACTGCAATGCCA 3221  
3181 TTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTAATGCCA 3240  
3222 CCTTACTTTGCCATGGAAAAATGCAGAGAAATCATAGATCAAAGCTTATACTCAAATAAA 3281  
3241 CCTTACTTTGCCATGGAAAAATGCAGAGAAATCAT- ----- 3274  
3282 CTTTCAAGGCTTTTGGCCATCTGCATATTTGGATGGACAAGCTGTTGTTGATATCCCTTAA 3341  
3275 -----AGGCTTTTGGCCATCTGCATATTTGGATGGACAAGCTGTTGTTGATATCCCTTAA 3329  
3342 TTTTATCATTTATTTTGTATGCTAGGAAGCTTTATGSCATTTTCATTTATGATATATAT 3401  
3330 TTTTATCATTTATTTTGTATGCTAGGAAGCTTTATGSCATTTTCATTTATGATATATAT 3389  
3402 TTTTATCATTTATTTTGTGCTGCTGCTGCTTTTGTGCTTTATTTGTTTATGTTTCAATCAGTT 3461  
3390 TTTTATCATTTATTTTGTGCTGCTGCTGCTTTTGTGCTTTATTTGCTTTATGTTTCAATCAGTT 3449  
3462 ATTCTGTTTCACTTATTTTGTGCTTTTCACTTTTACAGCTTTTAAAGAAATTTTAAATACCAAGAAATTT 3521  
3450 ATTCTGTTTCACTTATTTTGTGCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTT 3509  
3522 TGCTCATTTATCTATTTCTGTCAGAGCTTTGGCTTTGTTATGCAATCACTGAAATTAACCTTTC 3581  
3510 TGCTCATTTATCTATTTCTGTCAGAGCTTTGGCTTTGTTATGCAATCACTGAAATTAACCTTTC 3569  
3582 TTTATGGGATACAAATTTGCAACTTTCTTCAATTTATGCTTTTGTATCATCATTTCCAAATC 3641  
3570 TTTATGGGATACAAATTTGCAACTTTCTTCAATTTATGCTTTTGTATCATCATTTCCAAATC 3629  
3642 TATCCACTTTCTAGTTGCTGATTTCTTTCATAAAGATTTCTTGGAAAGATTTGACGAAA 3701  
3630 TATCCACTTTCTAGTTGCTGATTTCTTTCATAAAGATTTCTTGGAAAGATTTGACGAAA 3689  
3702 AATGTTGACACCTTATTAATCCATGGGATAGCTTTTCAAGTCTTTATATATGCTTTACCTG 3761  
3690 AATGTTGACACCTTATTAATCCATGGGATAGCTTTTCAAGTCTTTATATATGCTTTACCTG 3749  
3762 CAGTGTGCTAGTGGATTTTCTTCTTACAAATCATATGAGAAAAAATATGGAGGCAGATCA 3821  
3750 CAGTGTGCTAGTGGATTTTCTTCTTACAAATCATATGAGAAAAAATATGGAGGCAGATCA 3809  
3822 ATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAAGTCTTAAATATAGGAAGCTTCCA 3881  
3810 ATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAAGTCTTAAATATAGGAAGCTTCCA 3869  
3882 GAACCCACAGACAAATGAGGATGAGATGAGATGCAAGCTTCAAGAGCTTAAAGGTCAAA 3941  
3870 GAACCCACAGACAAATGAGGATGAGATGAGATGCAAGCTTCAAGAGCTTAAAGGTCAAA 3929

```
QY 3942 GAGCTGATGGTGGCAGTGTGTCAGAGAAACCATCATATATGTCAGCAATTTGCAT 4001
Db 3930 GAGCTGATGGTGGCAGTGTGTCAGAGAAACCATCATATATGTCAGCAATTTGCAT 3989
QY 4002 AAAGAAATATGATGACAGAAAGATTTCTTTTCAAGAAAAGTAAAGAAAGTGGCAACT 4061
Db 3990 AAAGAAATATGATGACAGAAAGATTTCTTTTCAAGAAAAGTAAAGAAAGTGGCAACT 4049
QY 4062 AAATACATCTCTTCTGTGTGAAAAGAGAGATCTTAGGACATATTTGGTCCAAATGGT 4121
Db 4050 AAATACATCTCTTCTGTGTGAAAAGAGAGATCTTAGGACATATTTGGTCCAAATGGT 4109
QY 4122 GCTGGCAAAAGCACAATTAATTAATCTCTGTGTGATATTTGAACCAACTTCAGGCCAG 4181
Db 4110 GCTGGCAAAAGCACAATTAATTAATCTCTGTGTGATATTTGAACCAACTTCAGGCCAG 4169
QY 4182 GTATTTTAAAGAGATTAATTTCTTCAGAGACAAGTGAAGATGATTCATCTGAAGTGTATG 4241
Db 4170 GTATTTTAAAGAGATTAATTTCTTCAGAGACAAGTGAAGATGATTCATCTGAAGTGTATG 4229
QY 4242 GGTACTGTCTCAGATAAACCTTTTGTGGCCAGATACCTACATTCGAGGAACATTTTGA 4301
Db 4230 GGTACTGTCTCAGATAAACCTTTTGTGGCCAGATACCTACATTCGAGGAACATTTTGA 4289
QY 4302 ATTTATGGAGCTGTCAAAGGAATGATGCAAGTGCATGAAAGAGTCAATAGTCGAATA 4361
Db 4290 ATTTATGGAGCTGTCAAAGGAATGATGCAAGTGCATGAAAGAGTCAATAGTCGAATA 4349
QY 4362 ACACATGCACTGTATTAAGAAACATCTTCAGAGACTGTAAAGAACTACCTGCAGGA 4421
Db 4350 ACACATGCACTGTATTAAGAAACATCTTCAGAGACTGTAAAGAACTACCTGCAGGA 4409
QY 4422 ATCAACAGAAAGTGTGTTTGTCTTAAGTATGCTAGGAATTCCTCAGATTAATTTGCTA 4481
Db 4410 ATCAACAGAAAGTGTGTTTGTCTTAAGTATGCTAGGAATTCCTCAGATTAATTTGCTA 4469
QY 4482 GATGAACCATCTACAGTATGATCCAAAGCCAAACAGACATCTGGCGAGCAATTCGA 4541
Db 4470 GATGAACCATCTACAGTATGATCCAAAGCCAAACAGACATCTGGCGAGCAATTCGA 4529
QY 4542 ACTGCATTTAAACAGAAAGCGGCTGCTATTCTGACCACTCACTATATGAGAGGCA 4601
Db 4530 ACTGCATTTAAACAGAAAGCGGCTGCTATTCTGACCACTCACTATATGAGAGGCA 4589
QY 4602 GAGGCTGTCTGTATCGATAGCTATCATGCTGTCTGGCAGTTAAGATGATCGAACA 4661
Db 4590 GAGGCTGTCTGTATCGATAGCTATCATGCTGTCTGGCAGTTAAGATGATCGAACA 4649
QY 4662 GTACAACATCTAAAGAGTAAATTTGAAAGGCTACTTTTGGAAATTAATTTGAAGAC 4721
Db 4650 GTACAACATCTAAAGAGTAAATTTGAAAGGCTACTTTTGGAAATTAATTTGAAGAC 4709
QY 4722 TGGATAGAAAACCTAGAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAAT 4781
Db 4710 TGGATAGAAAACCTAGAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAAT 4769
QY 4782 GCAAGCGCTCAGAAAGTGTCTCTATTCTGCTTATAAATTCCTAAGGAAGATGT 4841
Db 4770 GCAAGCGCTCAGAAAGTGTCTCTATTCTGCTTATAAATTCCTAAGGAAGATGT 4829
QY 4842 CAGTCCCTTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTCCCAATGAA 4901
Db 4830 CAGTCCCTTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTCCCAATGAA 4889
QY 4902 GAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAGACAA 4961
Db 4890 GAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAGACAA 4949
QY 4962 GAGGAGGAAGATTAATAGTCTGTGAACTTTTAAACAGCACACTTTTGTGGGAACGACACAA 5021
Db 4950 GAGGAGGAAGATTAATAGTCTGTGAACTTTTAAACAGCACACTTTTGTGGGAACGACACAA 5009
```

```
QY 5022 GAAGATAGAGTAGTATTTTGAATTTTGTATTGTTGGTCTGCTTACTGGGACTTCTTTCTT 5081
Db 5010 GAAGATAGAGTAGTATTTTGAATTTTGTATTGTTGGTCTGCTTACTGGGACTTCTTTCTT 5069
QY 5082 TTTCACTTAATTTTAACTTTGGTTTAAAAGTTTATTTTATTTGGAATGTTAACTGGAGAACC 5141
Db 5070 TTTCACTTAATTTTAACTTTGGTTTAAAAGTTTATTTTATTTGGAATGTTAACTGGAGAACC 5129
QY 5142 AAGAACGCACTTGAATTTTCTAAGCTCTCTTAATTTGAAATGCTGTGGTGTGTGTTTGG 5201
Db 5130 AAGAACGCACTTGAATTTTCTAAGCTCTCTTAATTTGAAATGCTGTGGTGTGTGTTTGG 5189
QY 5202 CTTTCTTTTAAATPAAAACCTATGTATTAATTAAGTGAAGCTGCATGTTTGTATGAAGTAT 5261
Db 5190 CTTTCTTTTAAATPAAAACCTATGTATTAATTAAGTGAAGCTGCATGTTTGTATGAAGTAT 5249
QY 5262 ATTGAACATATATGTTTGTATGTCATCTTTTCCACCATTCAGAAAACAGTCTTCTGAAT 5321
Db 5250 ATTGAACATATATGTTTGTATGTCATCTTTTCCACCATTCAGAAAACAGTCTTCTGAAT 5309
QY 5322 TGTGATTTAAAGGAATTTGTAATAGATAAGTTATTTTAAAGTATCTTTAAAGTTTATGC 5381
Db 5310 TGTGATTTAAAGGAATTTGTAATAGATAAGTTATTTTAAAGTATCTTTAAAGTTTATGC 5369
QY 5382 CATCTTCTTTAAATPAAAACCTATGTAATGTTCCAAATCTAATAAATAAATAATTAATTAAT 5441
Db 5370 CATCTTCTTTAAATPAAAACCTATGTAATGTTCCAAATCTAATAAATAAATAATTAATTAAT 5429
QY 5442 GCATAGAAAAGATACATAAAGCAATGTGAAAGTT 5475
Db 5430 GCATAGAAAAGATACATAAAGCAATGTGAAAGTT 5463

RESULT 4
AAD33648
ID AAD33648 standard; cdna; 6369 BP.
XX
XX AAD33648;
XX AC AC
XX DT 01-JUL-2002 (first entry)
XX Human TRICH-3 cdna.
XX
XX Human; transporter and ion channel; TRICH-3; transport disorder; angina;
XX amyotrophic lateral sclerosis; cystic fibrosis; neuromuscular disorder;
XX cardiac disorder; polymyositis; diabetes; neurological disorder; cancer;
XX depression; schizophrenia; anaemia; Wilson's disease; Cushing's disease;
XX cell proliferated disorder; infertility; arteriosclerosis; gene therapy;
XX Alzheimer's disease; Parkinson's disease; Huntington's disease; allergy;
XX myasthenia gravis; multiple sclerosis; metabolic disorder; hypertension;
XX acquired immune deficiency syndrome; immunological disorder; scleroderma;
XX endocrine disorder; autoimmune thyroiditis; rheumatoid arthritis; goitre;
XX cardiac myopathy; amnesia; toxic myopathy; Addison's disease; infection;
XX epilepsy; mental disorder; myocarditis; Crohn's disease; grave's disease;
XX muscle disorder; stroke; dementia; anxiety; AIDS; asthma; cirrhosis;
XX gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1238..6166
XX /*tag= a
XX /product= "Human TRICH-3 protein"
XX
XX WO200212340-A2.
XX
XX 14-FEB-2002.
XX
XX 01-AUG-2001; 2001WO-US024217.
XX
XX 03-AUG-2000; 2000US-0223269P.
XX 10-AUG-2000; 2000US-0224456P.
XX 18-AUG-2000; 2000US-0226410P.
XX
```





QY 1357 TCTATCTTGATCAAGTCAITCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 1416  
 Db 2481 TCTATCTTGATCAAGTCAITCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 2540  
 QY 1417 TGAAGCCTTCATATTGGTCAAGAGCAAAAGAAATTTATGAGGATTCAGAGGCAATG 1476  
 Db 2541 TGAAGCCTTCATATTGGTCAAGAGCAAAAGAAATTTATGAGGATTCAGAGGCAATG 2600  
 QY 1477 TTAATGGAATATTAGTTTTAGTGAATATTAGAGCCAGTTTCTTCAGAAATTTGAGGAA 1536  
 Db 2601 TTAATGGAATATTAGTTTTAGTGAATATTAGAGCCAGTTTCTTCAGAAATTTGAGGAA 2660  
 QY 1537 ABAAGCCATAAGAAATAGTGGTATTTCAGAGACATACAGAAAGAGGGTGAATATGCG 1596  
 Db 2661 ABAAGCCATAAGAAATAGTGGTATTTCAGAGACATACAGAAAGAGGGTGAATATGCG 2720  
 QY 1597 AGGCTTTGGAATTTGTCTATTTGACATATATGAGGGTCAGATTACTGCTTTACTTGGCC 1656  
 Db 2721 AGGCTTTGGAATTTGTCTATTTGACATATATGAGGGTCAGATTACTGCTTTACTTGGCC 2780  
 QY 1657 ACAGTGGACAGAAAGAGTACATTGATGAATATCTTTGTGGACTCTGCCACCTCTCG 1716  
 Db 2781 ACAGTGGACAGAAAGAGTACATTGATGAATATCTTTGTGGACTCTGCCACCTCTCG 2840  
 QY 1717 ATGGGTTTGCAATCTATATATGACACAGAGTCTCAGAAATAGATGAATTTGAGGAA 1776  
 Db 2841 ATGGGTTTGCAATCTATATATGACACAGAGTCTCAGAAATAGATGAATTTGAGGAA 2900  
 QY 1777 GAAATATGATGGCAATTTGTCCAGTTAGATATACATTTGATGTTTTGACAGTAGAAG 1836  
 Db 2901 GAAATATGATGGCAATTTGTCCAGTTAGATATACATTTGATGTTTTGACAGTAGAAG 2960  
 QY 1837 AAAATTTATCAATTTTGGCTTCAATCAAAGGGATACCCCAACAATATAATACAGGAA 1896  
 Db 2961 AAAATTTATCAATTTTGGCTTCAATCAAAGGGATACCCCAACAATATAATACAGGAA 3020  
 QY 1897 TGCAGAGGTTTTACTAGATTTAGACATGCAGACTATCAAGATCAACCAAGCTAAATAAT 1956  
 Db 3021 TGCAGAGGTTTTACTAGATTTAGACATGCAGACTATCAAGATCAACCAAGCTAAATAAT 3080  
 QY 1957 TAGTGGTGGTCAAAAGAGAGCTGTCAATAGGAATTCCTCTTGGGACCCCAAGA 2016  
 Db 3081 TAGTGGTGGTCAAAAGAGAGCTGTCAATAGGAATTCCTCTTGGGACCCCAAGA 3140  
 QY 2017 TACTGCTGTAGATGAACCAACAGCTGGAATGACCCCTGTTCTCGACATATTGTATGGA 2076  
 Db 3141 TACTGCTGTAGATGAACCAACAGCTGGAATGACCCCTGTTCTCGACATATTGTATGGA 3200  
 QY 2077 ATCTTTTAAATACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTCAAGGATG 2136  
 Db 3201 ATCTTTTAAATACAGAAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTCAAGGATG 3260  
 QY 2137 AAGCTGACATTTCTGAGATAGAAAGCTGTGATATCAAGGAATGCTGAAATGTTG 2196  
 Db 3261 AAGCTGACATTTCTGAGATAGAAAGCTGTGATATCAAGGAATGCTGAAATGTTG 3320  
 QY 2197 GTTCTTCAATGTTCTTCAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
 Db 3321 GTTCTTCAATGTTCTTCAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAG 3380  
 QY 2257 ACAATATTTGCGCACAGAAATCTTTCTTCACTGGTTAAACAAACATATACCTGGAGCTA 2316  
 Db 3381 ACAATATTTGCGCACAGAAATCTTTCTTCACTGGTTAAACAAACATATACCTGGAGCTA 3440  
 QY 2317 CTTTATACAGAAATGACCAACAACTTTGTGTATAGCTTGGCTTTCAAGGACATGGACA 2376  
 Db 3441 CTTTATACAGAAATGACCAACAACTTTGTGTATAGCTTGGCTTTCAAGGACATGGACA 3500  
 QY 2377 AATTTTCAGGTTTGTGTTTCTGGCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTATG 2436  
 Db 3501 AATTTTCAGGTTTGTGTTTCTGGCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTATG 3560

QY 2437 GTGTTTCCATGACGACTTTGGAAGACGTATTTTTAAAGCTAGAGTTGAAGCAGAAATG 2496  
 Db 3561 GTGTTTCCATGACGACTTTGGAAGACGTATTTTTAAAGCTAGAGTTGAAGCAGAAATG 3620  
 QY 2497 ACCAAGCAGATTTAGTGTATTTACTCAGCAGCCACTGGAGGAGAAATCGATTTCAAAAT 2556  
 Db 3621 ACCAAGCAGATTTAGTGTATTTACTCAGCAGCCACTGGAGGAGAAATCGATTTCAAAAT 3680  
 QY 2557 CTTTGTATGAAATGGAAACAGAGCTTACTTATTTCTGTAACCAAGGCTTCTCTAGTGA 2616  
 Db 3681 CTTTGTATGAAATGGAAACAGAGCTTACTTATTTCTGTAACCAAGGCTTCTCTAGTGA 3740  
 QY 2617 GCACCATGAGCCTTTGGAAACACAGATGTATCAATAGCAAAAGTTTCATTTCTTTACCT 2676  
 Db 3741 GCACCATGAGCCTTTGGAAACACAGATGTATCAATAGCAAAAGTTTCATTTCTTTACCT 3800  
 QY 2677 TGAACCTGAAAGTAAATCAGTGAGATCAAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG 2736  
 Db 3801 TGAACCTGAAAGTAAATCAGTGAGATCAAGTGTGCTTCTGCTTTTAAATTTTTTTCACAG 3860  
 QY 2737 TTCAGATTTTATGTTTTGTTTCATCAGCTTTTAAATGCTGTGTTTCCATCAAAAC 2796  
 Db 3861 TTCAGATTTTATGTTTTGTTTCATCAGCTTTTAAATGCTGTGTTTCCATCAAAAC 3920  
 QY 2797 TTGTTCCAGACTTATATTTTCTAAACCTTGAGACAAACCAATAATAACAAACAAAGTC 2856  
 Db 3921 TTGTTCCAGACTTATATTTTCTAAACCTTGAGACAAACCAATAATAACAAACAAAGTC 3980  
 QY 2857 TGCTTTTCAAAATTCGCTGACTCAGATATCAGTATCTTATAGCTTTTTCACAGGCC 2916  
 Db 3981 TGCTTTTCAAAATTCGCTGACTCAGATATCAGTATCTTATAGCTTTTTCACAGGCC 4040  
 QY 2917 AGAACATAATGTTGATGATGATTAATGACAGTATGATCCGTGGCTCCCATAGTG 2976  
 Db 4041 AGAACATAATGTTGATGATGATTAATGACAGTATGATCCGTGGCTCCCATAGTG 4100  
 QY 2977 CGGCTTTAAATGTGATGCAITTCAGAAAGGACTATGTTTTCAGAGCTGTTTTCACAGGTA 3036  
 Db 4101 CGGCTTTAAATGTGATGCAITTCAGAAAGGACTATGTTTTCAGAGCTGTTTTCACAGGTA 4160  
 QY 3037 CTATGGTTTATCTTTACCTATATAGTGAATATCAATAGTAATCTACTATCTTTATCATTT 3096  
 Db 4161 CTATGGTTTATCTTTACCTATATAGTGAATATCAATAGTAATCTACTATCTTTATCATTT 4220  
 QY 3097 TAAATGATGATGAAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAGAAATTTACTGATA 3156  
 Db 4221 TAAATGATGATGAAACCAATCCAGATCTGGAGTACCCCATTTCTTTCAGAAATTTACTGATA 4280  
 QY 3157 TAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGAATCATTTGTTACTGCAA 3216  
 Db 4281 TAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGAATCATTTGTTACTGCAA 4340  
 QY 3217 TGGCACTTACTTTGGCAATGGAAATGCAAGAAATCATAGATCAAGCTTATCTCAAC 3276  
 Db 4341 TGGCACTTACTTTGGCAATGGAAATGCAAGAAATCATAGATCAAGCTTATCTCAAC 4400  
 QY 3277 TTAACCTTTCAGCTCTTTGCTGCTGCTGCTTTTGGCTTATGTTGATATCC 3336  
 Db 4401 TTAACCTTTCAGCTCTTTGCTGCTGCTGCTTTTGGCTTATGTTGATATCC 4460  
 QY 3337 CCTTATTTTATCATCTTTTATTTGATGCTAGGAAAGCTTTATTTGGCAATTTTATTTGGAT 3396  
 Db 4461 CCTTATTTTATCATCTTTTATTTGATGCTAGGAAAGCTTTATTTGGCAATTTTATTTGGAT 4520  
 QY 3397 TATATTTTATCTGTAAGTTCCTTTCGCTGCTGCTTTTGGCTTATGTTTATGTTCCAT 3456  
 Db 4521 TATATTTTATCTGTAAGTTCCTTTCGCTGCTGCTTTTGGCTTATGTTTATGTTCCAT 4580  
 QY 3457 CAGTATTTCTGTTCACTTATTTGCTTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAAG 3516  
 Db 4581 CAGTATTTCTGTTCACTTATTTGCTTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAAG 4640  
 QY 3517 AATTTTGGTCAATTTATCTATTTCTGTGGGAGCGTGGCTGTGATTTGCAATCACTGAAATAA 3576

Db 4641 AATTTTGGTCATTTATCTATTTCTGTGGCAGCGTGGCTTGTATGTGAATCACTGAAATAA 4700  
 QY 3577 CTTTCTTTATGGGATACAAATTCGAATTAATCTTTCATTATATGCTTTTGTATCATCATTC 3635  
 Db 4701 CTTTCTTTATGGGATACAAATTCGAATTAATCTTTCATTATATGCTTTTGTATCATCATTC 4760  
 QY 3637 CAATCTATCCACTCTAGTGTGCTGATTTCTTTCATTAAGATTTCTTGGGAAGATGTAC 3695  
 Db 4761 CAATCTATCCACTCTAGTGTGCTGATTTCTTTCATTAAGATTTCTTGGGAAGATGTAC 4820  
 QY 3697 GAAAAATGTGGACACCTATAATCCATGGGATAGGCTTTCACTAGTGTGTATATCGCCTT 3756  
 Db 4821 GAAAAATGTGGACACCTATAATCCATGGGATAGGCTTTCACTAGTGTGTATATCGCCTT 4880  
 QY 3757 ACCTGAGTGTGTACGTGTGTATTTCTCTTCAATATCTATGAGAAAAAATATGAGGCA 3815  
 Db 4881 ACCTGAGTGTGTACGTGTGTATTTCTCTTCAATATCTATGAGAAAAAATATGAGGCA 4940  
 QY 3817 GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTCAAGGAAGTCTAAAAATAGGAAGC 3875  
 Db 4941 GATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTCAAGGAAGTCTAAAAATAGGAAGC 5000  
 QY 3877 TTCAGAACACACAGACAAATAGGATGAAGATGTAAGATGTAAGATGTAAGATGTAAGG 3935  
 Db 5001 TTCAGAACACACAGACAAATAGGATGAAGATGTAAGATGTAAGATGTAAGATGTAAGG 5060  
 QY 3937 TCAACAGCTGTATGGGTTCAGTGTGTGAGGAGAAACCTATCCATTATGGTCAGCAATT 3995  
 Db 5061 TCAACAGCTGTATGGGTTCAGTGTGTGAGGAGAAACCTATCCATTATGGTCAGCAATT 5120  
 QY 3997 TGCATAAAGAATATGATGACAGAAAGATTTCTTCTTTCAAGAAAGTAAAGAAAGTGG 4055  
 Db 5121 TGCATAAAGAATATGATGACAGAAAGATTTCTTCTTTCAAGAAAGTAAAGAAAGTGG 5180  
 QY 4057 CAACATAATACATCTTCTTCTGTGAAAAGAGGAGATCTTAGGACTATTTGGTCCAA 4115  
 Db 5181 CAACATAATACATCTTCTTCTGTGAAAAGAGGAGATCTTAGGACTATTTGGTCCAA 5240  
 QY 4117 ATGGTCTGGCAAAAGCAAAATTAATTAATTTCTGGTGGTATTTGAACCACTTCAG 4175  
 Db 5241 ATGGTCTGGCAAAAGCAAAATTAATTAATTTCTGGTGGTATTTGAACCACTTCAG 5300  
 QY 4177 GCAGGTATTTTAGAGATTAATTTCTTCAGACAAAGTGAAGATGATGATTCACCTGAAGT 4235  
 Db 5301 GCAGGTATTTTAGAGATTAATTTCTTCAGACAAAGTGAAGATGATGATTCACCTGAAGT 5360  
 QY 4237 GTATGGTGTACTGTCTCAGATAAACCCCTTTGTGGCCAGATACATTTGACAGGAACATT 4295  
 Db 5361 GTATGGTGTACTGTCTCAGATAAACCCCTTTGTGGCCAGATACATTTGACAGGAACATT 5420  
 QY 4297 TTGAATTTATGAGCTGTCAAAGGAATGAGTGCAGATGACATGAAGAAAGTCAATAAGTC 4355  
 Db 5421 TTGAATTTATGAGCTGTCAAAGGAATGAGTGCAGATGACATGAAGAAAGTCAATAAGTC 5480  
 QY 4357 GAATACACATGCACTGTGATTTAAAGACATCTTCAGAGACTCTTAAGAACTACCTG 4415  
 Db 5481 GAATACACATGCACTGTGATTTAAAGACATCTTCAGAGACTCTTAAGAACTACCTG 5540  
 QY 4417 CAGGAATCAAAAGAAAGTGTGTTTTTCTTAAGTATGCTAGGGAATCTCTCAGATTACTT 4475  
 Db 5541 CAGGAATCAAAAGAAAGTGTGTTTTTCTTAAGTATGCTAGGGAATCTCTCAGATTACTT 5600  
 QY 4477 TGTAGATGAACCATCTACAGGTATGGATCCCAAGCCCAACAGACATGTGGCCAGCAA 4535  
 Db 5601 TGTAGATGAACCATCTACAGGTATGGATCCCAAGCCCAACAGACATGTGGCCAGCAA 5660  
 QY 4537 TTCGAAGTCAATTTAAAAACAGAAAGCGGCTGCTATCTGACCACTCACTATATGAGG 4595  
 Db 5661 TTCGAAGTCAATTTAAAAACAGAAAGCGGCTGCTATCTGACCACTCACTATATGAGG 5720  
 QY 4597 AGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGGAGTTAGATGTATCG 4655

Db 5721 AGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGCAGTTAAGATGTATCG 5780  
 QY 4657 GAACAGTACAACATCTAAAGAGTAAATTTGGAAGAGCTACTTTTTGGAATTAATAATTGA 4716  
 Db 5781 GAACAGTACAACATCTAAAGAGTAAATTTGGAAGAGCTACTTTTTGGAATTAATAATTGA 5840  
 QY 4717 AGGACTGATAGAAAACCTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCC 4776  
 Db 5841 AGGACTGATAGAAAACCTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCC 5900  
 QY 4777 CAATGCAAGCCGTCAGAAAGTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAG 4836  
 Db 5901 CAATGCAAGCCGTCAGAAAGTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAG 5960  
 QY 4837 ATGTTCACTGCTCTTCAATCTTTTAACTGGAAGAGTAAACATGCTTTTGCCA 4896  
 Db 5961 ATGTTCACTGCTCTTCAATCTTTTAACTGGAAGAGTAAACATGCTTTTGCCA 6020  
 QY 4897 TTGAAGATATAGCTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAG 4956  
 Db 6021 TTGAAGATATAGCTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAAG 6080  
 QY 4957 AACAGAGGAGAGATATAGTTGTGGAACCTTTAAACAGCACACTTTGGTGGGAACGAA 5016  
 Db 6081 AACAGAGGAGAGATATAGTTGTGGAACCTTTAAACAGCACACTTTGGTGGGAACGAA 6140  
 QY 5017 CACAAGAGATAGAGTAGTATTTGAATTTGTATTTGTGGTCTGCTTACTGGGACTTCT 5076  
 Db 6141 CACAAGAGATAGAGTAGTATTTGAATTTGTATTTGTGGTCTGCTTACTGGGACTTCT 6200  
 QY 5077 TTTCTTTTCACTTAATTTTAACTTTGGTTTAAAAAGTTTTTTTATTTGAATGGAATGGAATGGA 5136  
 Db 6201 TTTCTTTTCACTTAATTTTAACTTTGGTTTAAAAAGTTTTTTTATTTGAATGGAATGGAATGGA 6260  
 QY 5137 GAAACAGAGAGCAGCTTGAATTTTCTAGACTCTTAAATGGAATGCTGCTGTTGTG 5196  
 Db 6261 GAAACAGAGAGCAGCTTGAATTTTCTAGACTCTTAAATGGAATGCTGCTGTTGTG 6320  
 QY 5197 TTTTGTCTTTCTTTAAATAAAACGATGATGATAAATTAAGTAA 5238  
 Db 6321 TTTTGTCTTTCTTTAAATAAAACGATGATGATAAATTAAGTAA 6362

RESULT 5  
 AAD37620  
 ID AAD37620 standard; cdna; 5262 BP.  
 XX  
 AC AAD37620;  
 XX  
 DT 10-SEP-2002 (first entry)  
 XX  
 DE Human transporter protein cdna #3.  
 DE  
 KW Human; novel human protein; NHP; transporter protein; mental disorder;  
 KW cancer; gene therapy; drug screening; nutraceutical application;  
 KW cosmetic application; polymorphism; ss.  
 XX  
 OS Homo sapiens.  
 OS  
 FN WO200231147-A2.  
 XX  
 PD 18-APR-2002.  
 XX  
 PF 04-OCT-2001; 2001WO-US031113.  
 XX  
 PR 10-OCT-2000; 2000US-0239629P.  
 XX  
 PA (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Hu Y, Nepomnichy B;  
 XX  
 DR WPI; 2002-454552/48.  
 XX

PT Novel nucleic acid molecule encoding novel human proteins, useful for  
 PT therapeutic, diagnostic and pharmacogenomic applications.  
 XX Disclosure; Page 45-46; 46pp; English.  
 XX The present sequence is a cDNA coding for novel human protein (NHP),  
 CC human transporter protein. NHPs shares structural similarity with the  
 CC mammalian ATP-binding cassette (ABC) transporters and multidrug  
 CC resistance transporters. NHP polynucleotides are useful for the  
 CC therapeutic, diagnostic and pharmacogenomic applications. They are used  
 CC for detecting and treating mental disorders and cancers. They are also  
 CC used in gene therapy. NHP polypeptides are useful for diagnosis, drug  
 CC screening, clinical trial monitoring, treatment of diseases and  
 CC disorders, and cosmetic or nutraceutical applications  
 XX Sequence 5262 BP; 1615 A; 888 C; 1012 G; 1743 T; 0 U; 4 Other;  
 SQ  
 Query Match 91.6%; Score 5014.6; DB 6; Length 5262;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 5014; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 QY 97 AGGTTTATTCAGAAAACATGCTCCACTGCAATTAGGGAGTAGGAGTTGGAGACAGACCA 156  
 DB 238 AGGTTTATTCAGAAAACATGCTCCACTGCAATTAGGGAGTAGGAGTTGGAGACAGACCA 297  
 QY 157 GAACACTTCTACTGAAGAATTAATTAATTAATGAGAACCAAAAGAGTAGTTGAGG 216  
 DB 298 GAACACTTCTACTGAAGAATTAATTAATTAATGAGAACCAAAAGAGTAGTTGAGG 357  
 QY 217 AAATCTCTTTTCCACTATTTTATTTTGGTTAATTAATTAATGAGTAGGATGATCCAA 276  
 DB 358 AAATCTCTTTTCCACTATTTTATTTTGGTTAATTAATTAATGAGTAGGATGATCCAA 417  
 QY 277 ATAGAAATATGAGAAAGTGCCCTAATATAGAACTCAATCTTATGAGCAAGTTACTCTTT 336  
 DB 418 ATAGAAATATGAGAAAGTGCCCTAATATAGAACTCAATCTTATGAGCAAGTTACTCTTT 477  
 QY 337 CTAATCTAATCTTGATATATCTCCAGTCACTAATTAATTAATGAGTAGGATGATCCAA 396  
 DB 478 CTAATCTAATCTTGATATATCTCCAGTCACTAATTAATTAATGAGTAGGATGATCCAA 537  
 QY 397 TGTCTACTGATCATCTACCTGATGTCTAATTAATTAATGAGTAGGATGATCCAA 456  
 DB 538 TGTCTACTGATCATCTACCTGATGTCTAATTAATTAATGAGTAGGATGATCCAA 597  
 QY 457 TGTAAACATCCAGTCTCTTAAGCCGAGCAACTTTGAGTGTGCTTTCAGAGACTCCA 516  
 DB 598 TGTAAACATCCAGTCTCTTAAGCCGAGCAACTTTGAGTGTGCTTTCAGAGACTCCA 657  
 QY 517 TGTCTCTATGAACCTTCGTTTTTCTCTGATATGATCCAGTATCTTCTATTTATATGAGTT 576  
 DB 658 TGTCTCTATGAACCTTCGTTTTTCTCTGATATGATCCAGTATCTTCTATTTATATGAGTT 717  
 QY 577 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTTGCTCCTCAGTTTCAG 636  
 DB 718 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTTGCTCCTCAGTTTCAG 777  
 QY 637 TTTTCAAGCATCCATAGATGCTGCCATTAATCAGTTGAAGCACTGTTCTCTTTTGA 696  
 DB 778 TTTTCAAGCATCCATAGATGCTGCCATTAATCAGTTGAAGCACTGTTCTCTTTTGA 837  
 QY 697 AGGAGCTGGAGTCAACTAAAGCTGTTTATTAATGAGGAACTGCTGTTGAGAAATAGATA 756  
 DB 838 AGGAGCTGGAGTCAACTAAAGCTGTTTATTAATGAGGAACTGCTGTTGAGAAATAGATA 897  
 QY 757 CTTTCCCGGAGGTAATTTAATATACCTAGTTATAGCACTTTTCACTTTTGTGATACT 816  
 DB 898 CTTTCCCGGAGGTAATTTAATATACCTAGTTATAGCACTTTTCACTTTTGTGATACT 957  
 QY 817 TTTTGGCAATTCATATCGTAGCAGAAAAGAAAAATAAAGAAATTTTAAAGATAA 876  
 DB 958 TTTTGGCAATTCATATCGTAGCAGAAAAGAAAAATAAAGAAATTTTAAAGATAA 1017

QY 877 TGGGACTTCATGATGATGCTGCTTTTGGCTTTCTCTATATACAAGTTTAAATTT 936  
 DB 1018 TGGGACTTCATGATGATGCTGCTTTTGGCTTTCTCTATATACAAGTTTAAATTT 1077  
 QY 937 TTTCTATGCTCCCTTTTATGAGGAGTCAATGCGACAGCTTCTTTTGTATTTTCTCAAAGTA 996  
 DB 1078 TTTCTATGCTCCCTTTTATGAGGAGTCAATGCGACAGCTTCTTTTGTATTTTCTCAAAGTA 1137  
 QY 997 GCAGCAATGATATTTCTGCTTTTCTTTTATGAGTATCATCTGATTTTGTGCTT 1056  
 DB 1138 GCAGCAATGATATTTCTGCTTTTCTTTTATGAGTATCATCTGATTTTGTGCTT 1197  
 QY 1057 TAAATGTCACACCTCTTTTAAATAAATAAACAATGCGGAATGAGTTTGTGTA 1116  
 DB 1198 TAAATGTCACACCTCTTTTAAATAAATAAACAATGCGGAATGAGTTTGTGTA 1257  
 QY 1117 CTGTGCTTTTGGATTTATTTGGCTTTATGATATCTCTATAGAAAGTTTCCCAATCGT 1176  
 DB 1258 CTGTGCTTTTGGATTTATTTGGCTTTATGATATCTCTATAGAAAGTTTCCCAATCGT 1317  
 QY 1177 TAGTGTGGCTTTTCACTGCTTTCTGTCACGTACTGTTTGTGATTTGATTTGACAGGTCA 1236  
 DB 1318 TAGTGTGGCTTTTCACTGCTTTCTGTCACGTACTGTTTGTGATTTGATTTGACAGGTCA 1377  
 QY 1237 TGCATTTTAGAAGATTTTAAATGAGGTCCTTCAATTTTCAAAATTTGACTGAGGCCATATC 1296  
 DB 1378 TGCATTTTAGAAGATTTTAAATGAGGTCCTTCAATTTTCAAAATTTGACTGAGGCCATATC 1437  
 QY 1297 CTCTAATTTATTAACAATTTATCATGCTCACACTTAAATAGTATTTCTATGCTCTTGGCTG 1356  
 DB 1438 CTCTAATTTATTAACAATTTATCATGCTCACACTTAAATAGTATTTCTATGCTCTTGGCTG 1497  
 QY 1357 TCTATCTTGTCAAGTCACTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 1416  
 DB 1498 TCTATCTTGTCAAGTCACTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTC 1557  
 QY 1417 TGAAGCCTTCATATTTGCTCAAGAGCAAAAGAAATTTATGAGGAGTTTATCAGAGGGCAATG 1476  
 DB 1558 TGAAGCCTTCATATTTGCTCAAGAGCAAAAGAAATTTATGAGGAGTTTATCAGAGGGCAATG 1617  
 QY 1477 TTAATGGAATATTTAGTTTGTGAAATTTATGAGCAGTTTCTTCAAGTTTGTAGGA 1536  
 DB 1618 TTAATGGAATATTTAGTTTGTGAAATTTATGAGCAGTTTCTTCAAGTTTGTAGGA 1677  
 QY 1537 AAGAAGCCATTAAGAAATTTAGTGTATTTTCAAGAGACATACAGAAAGAGGGTGAATGTGG 1596  
 DB 1578 AAGAAGCCATTAAGAAATTTAGTGTATTTTCAAGAGACATACAGAAAGAGGGTGAATGTGG 1737  
 QY 1597 AGGCTTTGAGAAATTTGTCATTTTGAATATATGAGGCTCAGATTACTGCTTACTTTGGCC 1656  
 DB 1738 AGGCTTTGAGAAATTTGTCATTTTGAATATATGAGGCTCAGATTACTGCTTACTTTGGCC 1797  
 QY 1657 ACAGTGGAAACAGAAAGAGTACATTGATGAATTTCTTTGTGACTCTGCCACCTTTCTG 1716  
 DB 1798 ACAGTGGAAACAGAAAGAGTACATTGATGAATTTCTTTGTGACTCTGCCACCTTTCTG 1857  
 QY 1717 ATGGTTTGTCAATCTATATATGAGACAGAGTCTCAGAAATAGATGAATTTGTTGAAGCAA 1776  
 DB 1858 ATGGTTTGTCAATCTATATATGAGACAGAGTCTCAGAAATAGATGAATTTGTTGAAGCAA 1917  
 QY 1777 GAAATATGATTTGGCAATTTGTCACAGTTAGATATACACTTTGATTTTGTGACAGTAGAAG 1836  
 DB 1918 GAAATATGATTTGGCAATTTGTCACAGTTAGATATACACTTTGATTTTGTGACAGTAGAAG 1977  
 QY 1837 AAAATTTATCAATTTTGGCTTCAATCAAAGGGGATACAGCCAAATATAATTAACAAGAG 1896  
 DB 1978 AAAATTTATCAATTTTGGCTTCAATCAAAGGGGATACAGCCAAATATAATTAACAAGAG 2037  
 QY 1897 TGCAGAGAGTTTACTAGATTTTAGACATGCAGACTATCAAGATACACCAAGCTTAAATAT 1956  
 DB 2038 TGCAGAGAGTTTACTAGATTTTAGACATGCAGACTATCAAGATACACCAAGCTTAAATAT 2097  
 QY 1957 TAAGTGTGCTCAAAAAGAAAGAGCTGTCTATTAGGAATTTGCTGTTTGGGAACCCCAAGA 2016

2098 TAAGTGGTGGTCAAAAGAAAGAGCTGTCAATAGGAATTCCTGTTCTTGGGAACCCAAAGA 2157  
2017 TACTGCTGTAGATGAACCAACAGCTGGAAATGGAGCCCTGTTCTCGACATATTCGTATGGA 2076  
2158 TACTGCTGTAGATGAACCAACAGCTGGAAATGGAGCCCTGTTCTCGACATATTCGTATGGA 2217  
2077 ATCTTTTAAATACAGAAAGCAATCGGGTGACAGTGTTCAGTACTCAATTCATGGATG 2136  
2218 ATCTTTTAAATACAGAAAGCAATCGGGTGACAGTGTTCAGTACTCAATTCATGGATG 2277  
2137 AGCTGACATTCCTCGAGATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTGTG 2196  
2278 AAGCTGACATTCCTCGAGATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTGTG 2337  
2197 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2256  
2338 GTTCTTCAATGTTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAG 2397  
2257 ACAATATTTGCGCACAGAACTCTTCTTCACTCGGTTAAACAACATATACCTGGAGCTA 2316  
2398 ACAATATTTGCGCACAGAACTCTTCTTCACTCGGTTAAACAACATATACCTGGAGCTA 2457  
2317 CTTTTATTACAGAGATGACCAACAACTTGTGTATAGCTTGCCTTTCAAGGACATGACA 2376  
2458 CTTTATTTACAGAGATGACCAACAACTTGTGTATAGCTTGCCTTTCAAGGACATGACA 2517  
2377 AATTTTCAGGTTTGTTCGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATG 2436  
2518 AATTTTCAGGTTTGTTCGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATG 2577  
2437 GTGTTTCAGTACAGCTTGTGAAGACGTATTTTAAAGCTAGAGTTGAGCAGAAATG 2496  
2578 GGGTTTCAGTACAGCTTGTGAAGACGTATTTTAAAGCTAGAGTTGAGCAGAAATG 2637  
2497 ACCAAGCAGATTTATAGTGTATTTACTCAGCAGCCACTGGAGGAAAGATGATTCAAAAT 2556  
2638 ACCAAGCAGATTTATAGTGTATTTACTCAGCAGCCACTGGAGGAAAGATGATTCAAAAT 2697  
2557 CTTTGTATGAATGGAACAGAGCTTACTTATCTTCTTGAACCAAGGCTTCTAGTGA 2616  
2698 CTTTGTATGAATGGAACAGAGCTTACTTATCTTCTTGAACCAAGGCTTCTAGTGA 2757  
2617 GCACCATGAGCCCTTTGGAAAACACAGATGTATACAATAGCAAAAGTTTCATTTCTTTACCT 2676  
2758 GCACCATGAGCCCTTTGGAAAACACAGATGTATACAATAGCAAAAGTTTCATTTCTTTACCT 2817  
2677 TGAAACGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTAATTTTTCACAG 2736  
2818 TGAAACGTGAAGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTAATTTTTCACAG 2877  
2737 TTCAGATTTTATGTTTGTGTTTCACTCACTCTTTTAAAAATGCTGTGTTCCCATCAAC 2796  
2878 TTCAGATTTTATGTTTGTGTTTCACTCACTCTTTTAAAAATGCTGTGTTCCCATCAAC 2937  
2797 TTGTTCCAGACTTATTTTCAAAACCTGGAGACAAACCAATATAACAAAACAAAGTC 2856  
2938 TTGTTCCAGACTTATTTTCAAAACCTGGAGACAAACCAATATAACAAAACAAAGTC 2997  
2857 TGCTTCTTCAAAATTCGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 2916  
2998 TGCTTCTTCAAAATTCGCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCC 3057  
2917 AGAACATATGCTGACATGATTAATGACAGTACTATGATCCGTGGCTCCCATAGTG 2976  
3058 AGAACATATGCTGACATGATTAATGACAGTACTATGATCCGTGGCTCCCATAGTG 3117  
2977 CGGCTTTAAATGATGATTCAGAAAAGGACTATGTTTTGCGAGCTGTTTTCAACAGTA 3036  
3118 CGGCTTTAAATGATGATTCAGAAAAGGACTATGTTTTGCGAGCTGTTTTCAACAGTA 3177  
3037 CTATGGTTTATCTTTTACTTATTTAGTGAATATCATTAGTAACCTACTATCTTTTATCAT 3096

3178 CTATGGTTTATCTTTTACCTATATTAAGTGAATATCATTAAGTAACTACTATCTTTATCAT 3237  
3097 TAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAGAAAATTTACTGATA 3156  
3238 TAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAGAAAATTTACTGATA 3297  
3157 TAGTTTTTAAAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTACTGCA 3216  
3298 TAGTTTTTAAAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTACTGCA 3357  
3217 TGGCACCCTTACCTTGGCATGGAAATGCAAGAGATCATTAAGATCAAGCTTATACCTCAAC 3276  
3358 TGGCACCCTTACCTTGGCATGGAAATGCAAGAGATCATTAAGATCAAGCTTATACCTCAAC 3417  
3277 TTTAAACCTTTTCAGCTCTTTTGGCCATCTGCATATTTGGATTTGGCAAGCTGTTGTGATCC 3336  
3418 TTTAAACCTTTTCAGCTCTTTTGGCCATCTGCATATTTGGATTTGGCAAGCTGTTGTGATCC 3477  
3337 CCTTATTTTATCATCTTATTTTGAATCTAGGAGCTTATGGCAATTTCAATATGGAT 3396  
3478 CCTTATTTTATCATCTTATTTTGAATCTAGGAGCTTATGGCAATTTCAATATGGAT 3537  
3397 TATATTTTATACCTGTAAAGTTCCTTGTCTGTGTTTTTGGCCTTATTTGGTATGTTCCAT 3456  
3538 TATATTTTATACCTGTAAAGTTCCTTGTCTGTGTTTTTGGCCTTATTTGGTATGTTCCAT 3597  
3457 CAGTATTCCTGTCACCTTATTTGCTTCTTCACTTTAAGAAAATTTTAAATACCAAAG 3516  
3598 CAGTATTCCTGTCACCTTATTTGCTTCTTCACTTTAAGAAAATTTTAAATACCAAAG 3657  
3517 AATTTTGTGTCATTTATCTATCTGCGCAGCTTGGCTGTATTTGCAATCACTGAATAA 3576  
3658 AATTTTGTGTCATTTATCTATCTGCGCAGCTTGGCTGTATTTGCAATCACTGAATAA 3717  
3577 CTTTCTTTATGGGATACAAATTTGCAATTTCTTCAATATGCTTTTGTATCATCATTC 3636  
3718 CTTTCTTTATGGGATACAAATTTGCAATTTCTTCAATATGCTTTTGTATCATCATTC 3777  
3637 CAATCTATCCACTTCTAGGTTGCTGATTTCTTCAATGAATTTCTTGGAAAGATGTAC 3696  
3778 CAATCTATCCACTTCTAGGTTGCTGATTTCTTCAATGAATTTCTTGGAAAGATGTAC 3837  
3697 GAAAAATGTGACACCTTAAATCCATGGGATAGGCTTTTCAGTGTGTTATATGCGCTT 3756  
3838 GAAAAATGTGACACCTTAAATCCATGGGATAGGCTTTTCAGTGTGTTATATGCGCTT 3897  
3757 ACCTGAGTGTGATCTGTGAGATTTTCTCTTACAATATCTAGAAAATATGGAGCA 3816  
3898 ACCTGAGTGTGATCTGTGAGATTTTCTCTTACAATATCTAGAAAATATGGAGCA 3957  
3817 GATCAATAAGAAAAGATCCCTTTTTCAGAAAACCTTTTCAACGAAGTCTAAAAATAGGAAGC 3876  
3958 GATCAATAAGAAAAGATCCCTTTTTCAGAAAACCTTTTCAACGAAGTCTAAAAATAGGAAGC 4017  
3877 TTCAGAAACCAACAGCAATGAGGATGAAAGATGAAAGTCAAGCTGAAAGCTAAAG 3936  
4018 TTCAGAAACCAACAGCAATGAGGATGAAAGTCAAGTCAAGCTGAAAGCTAAAG 4077  
3937 TCAGAGCTGTGAGGTTGCTGAGTGTGAGGAGAAACCATCCATTTATGCTGAGCAAT 3996  
4078 TCAGAGCTGTGAGGTTGCTGAGTGTGAGGAGAAACCATCCATTTATGCTGAGCAAT 4137  
3997 TGCAATAAGAAATATGATGACAAAGATTTTCTTTTCAAGAAAAGTAAAGAAAGTGG 4056  
4138 TGCAATAAGAAATATGATGACAAAGATTTTCTTTTCAAGAAAAGTAAAGAAAGTGG 4197  
4057 CAATCAATACATCTCTTCTGTGAAAAGAGGAGATCTTAGGACTATTTGGGTCCTAA 4116  
4198 CAATCAATACATCTCTTCTGTGAAAAGAGGAGATCTTAGGACTATTTGGGTCCTAA 4257  
4117 ATGGTCTGGCAAAAGCAAAATTTATTTATTTCTGTTGGTGTGATATGAACTTCTAG 4176  
4258 ATGGTCTGGCAAAAGCAAAATTTATTTATTTCTGTTGGTGTGATATGAACTTCTAG 4317

```
QY 4177 GCCAGGTATTTTAGGAGATTATCTTCAGAGACAAGTGAAGATGATGATTCATCTGAAGT 4236
XX
DQ 4318 GCCAGGTATTTTAGGAGATTATCTTCAGAGACAAGTGAAGATGATGATTCATCTGAAGT 4377
KW
KW 4237 GTATGGGTACTGCTCCTCAGATAAACCCCTTTGTCGCGAGATACATACATTCAGAGAACATT 4296
KW
OS 4378 GTATGGGTACTGCTCCTCAGATAAACCCCTTTGTCGCGAGATACATACATTCAGAGAACATT 4437
XX
XX 4297 TTGAAATTTATGGAGCTGTCAAGGAATGAGTCAAGTCAATGAAAGAGTCAATAGTC 4356
FH
FH 4438 TTGAAATTTATGGAGCTGTCAAGGAATGAGTCAAGTCAATGAAAGAGTCAATAGTC 4497
FT
FT 4357 GAATAACACATGCTATGATTTAAAGAAACATCTTCAGAGACTGTGTAAGAACTACCTG 4416
FT
FT 4498 GAATAACACATGCTATGATTTAAAGAAACATCTTCAGAGACTGTGTAAGAACTACCTG 4557
FT
QY 4417 CAGGAATCAACAGCAAGTGTGTTGCTCTTAAGTATGCTAGGGAATCCCTCAGATTACTT 4476
XX
DQ 4558 CAGGAATCAACAGCAAGTGTGTTGCTCTTAAGTATGCTAGGGAATCCCTCAGATTACTT 4617
XX
QY 4477 TCTAGATGAACCATCTACAGTATGATCCCAAGCCAAACACACATGTGGCGAGCA 4536
DQ 4618 TCTAGATGAACCATCTACAGTATGATCCCAAGCCAAACACACATGTGGCGAGCA 4577
XX
QY 4537 TCTGAACTGCAATTTAAAGAAAGAGCGGGCTCTTCTGACCACTCCTATATGAGG 4596
DQ 4678 TCTGAACTGCAATTTAAAGAAAGAGCGGGCTCTTCTGACCACTCCTATATGAGG 4737
XX
QY 4597 AGGCAGAGGCTGTCTGTGATCAGTATGCTATGCTGTGCGAGCTTAAGATGATCG 4656
DQ 4738 AGGCAGAGGCTGTCTGTGATCAGTATGCTATGCTGTGCGAGCTTAAGATGATCG 4797
XX
QY 4657 GAACAGTACACATCTAAGAGTAAATTTGAAAGAGCTACTTTTGGAAATTAATTA 4716
DQ 4798 GAACAGTACACATCTAAGAGTAAATTTGAAAGAGCTACTTTTGGAAATTAATTA 4857
XX
QY 4717 AGGACTGGATAGAAACCTAGAGTAGACCGCTTTCAAGAGAAATTCAGTATATTTCC 4776
DQ 4858 AGGACTGGATAGAAACCTAGAGTAGACCGCTTTCAAGAGAAATTCAGTATATTTCC 4917
XX
QY 4777 CAATGCAAGCGCTCAGAAAGTCTTCTTCTATTTTGGCTTATAAAATTCCTAAGAG 4836
DQ 4918 CAATGCAAGCGCTCAGAAAGTCTTCTTCTATTTTGGCTTATAAAATTCCTAAGAG 4977
XX
QY 4837 ATGTTTCAGTCCCTTTTCAAAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTGCCA 4896
DQ 4978 ATGTTTCAGTCCCTTTTCAAAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTGCCA 5037
XX
QY 4897 TTGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGCTTTTGTAGACTCACTAAG 4956
DQ 5038 TTGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGCTTTTGTAGACTCACTAAG 5097
XX
QY 4957 AACAGAGGAGGAGATATAGTTGTGAACTTTTAAACAGACACTTTGTTGGGAGCA 5016
DQ 5098 AACAGAGGAGGAGATATAGTTGTGAACTTTTAAACAGACACTTTGTTGGGAGCA 5157
XX
QY 5017 CACAGAGATAGATAGTATTTGAAATTTGATTTGCTGCTGCTTACTGGACATCT 5076
DQ 5158 CACAGAGATAGATAGTATTTGAAATTTGATTTGCTGCTGCTTACTGGACATCT 5217
XX
QY 5077 TTTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGATTTT 5117
DQ 5218 TTTCTTTTCACTTAATTTTAACTTTTGGTTTAAAGATTTT 5258
XX
```

## RESULT 6

AAD37618

ID AAD37618 standard; cDNA; 4929 BP.

XX AAD37618;

AC AAD37618;

XX 10-SEP-2002 (first entry)

```
XX
DE Human transporter protein cDNA #1.
XX
KW Human; novel human protein; NHP; transporter protein; mental disorder;
KW cancer; gene therapy; drug screening; nutraceutical application; gene;
KW cosmetic application; polymorphism; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 1..4929
FT /tag= a
FT /product= "Human transporter protein #1"
FT /transl_except= (pos:808..810, aa:Tyr)
FT /transl_except= (pos:2494..2496, aa:Ser)
FT /transl_except= (pos:2878..2880, aa:Val)
FT misc_feature 810
FT /tag= b
FT /note= "This degenerate base represents a polymorphic
FT site"
FT misc_feature 2494
FT /tag= c
FT /note= "This degenerate base represents a polymorphic
FT site"
FT misc_feature 2878
FT /tag= d
FT /note= "This degenerate base represents a polymorphic
FT site"
XX
PN WO200231147-A2.
XX
PD 18-APR-2002.
XX
PF 04-OCT-2001; 2001WO-US031113.
XX
PR 10-OCT-2000; 2000US-0239629P.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Hu Y, Nepomnichy B;
XX
XX WPI; 2002-454552/48.
XX P-PSDB; AAE23856.
XX
XX Novel nucleic acid molecule encoding novel human proteins, useful for
XX therapeutic, diagnostic and pharmacogenomic applications.
XX
XX Claim 1; Page 34-35; 46pp; English.
XX
XX The present sequence is a cDNA coding for novel human protein (NHP),
XX human transporter protein. NHPs shares structural similarity with the
XX mammalian ATP-binding cassette (ABC) transporters and multidrug
XX resistance transporters. NHP polynucleotides are useful for the
XX therapeutic, diagnostic and pharmacogenomic applications. They are used
XX for detecting and treating mental disorders and cancers. They are also
XX used in gene therapy. NHP polypeptides are useful for diagnosis, drug
XX screening, clinical trial monitoring, treatment of diseases and
XX disorders, and cosmetic or nutraceutical applications
XX
SQ Sequence 4929 BP; 1544 A; 827 C; 950 G; 1604 T; 0 U; 4 Other;
```

Query Match 89.9%; Score 4922.6; DB 6; Length 4929;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 4922; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 114 ATGTCACATGCAATAGGAGCTAGGAGTTTGGAGACAGACACACTTCTACTGAAG 173

DQ 1 ATGTCACATGCAATAGGAGCTAGGAGTTTGGAGACAGACACACTTCTACTGAAG 60

QY 174 AATTACTTAATTAATGCAAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 233

DQ 61 AATTACTTAATTAATGCAAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 120

QY 234 TTTTCTTTTATTTGGTTAAATTAATTAAGCATGATGCATCCAAATAAGATATGAAGAA 293  
DB 121 TTTTCTTTTATTTGGTTAAATTAATTAAGCATGATGCATCCAAATAAGATATGAAGAA 180  
QY 294 GTGCCATAATAGAACTCAATCTCTATGACCAAGTTTACTCTTTCTAATCTAATTTCTTGGGA 353  
DB 181 GTGCCATAATAGAACTCAATCTCTATGACCAAGTTTACTCTTTCTAATCTAATTTCTTGGGA 240  
QY 354 TATACCTCAGTGACTAATATATACAGCAGCAGCATCATGCAAGAAAGTGTCTACTGATCATCTA 413  
DB 241 TATACCTCAGTGACTAATATATACAGCAGCAGCATCATGCAAGAAAGTGTCTACTGATCATCTA 300  
QY 414 COTGATGTCAATAATTAAGAAATATACAAATGAAGAAAGAAATGTTAAACATCCAGTCTC 473  
DB 301 COTGATGTCAATAATTAAGAAATATACAAATGAAGAAAGAAATGTTAAACATCCAGTCTC 360  
QY 474 TCTAAGCCGAGCACTTTGTAGGTGTGTTTCAAGACTCCATGCTCTATGAACCTTCGT 533  
DB 361 TCTAAGCCGAGCACTTTGTAGGTGTGTTTCAAGACTCCATGCTCTATGAACCTTCGT 420  
QY 534 TTTTCTCTGATGATGATCCAGTATCTTCTATTTATATGGAATCAAGAGCTGGCTGTTC 593  
DB 421 TTTTCTCTGATGATGATCCAGTATCTTCTATTTATATGGAATCAAGAGCTGGCTGTTC 480  
QY 594 AAATCATGTGAGGCTGCTCAGTACGTGTCCTCAGGTTTCAAGTTTCAAGCATCCATA 653  
DB 481 AAATCATGTGAGGCTGCTCAGTACGTGTCCTCAGGTTTCAAGTTTCAAGCATCCATA 540  
QY 654 GATGCTGCCATTAATACAGTTGGAAGCAATGTTTCTCTTGGAGAGGAGCTGGAGTCAACT 713  
DB 541 GATGCTGCCATTAATACAGTTGGAAGCAATGTTTCTCTTGGAGAGGAGCTGGAGTCAACT 600  
QY 714 AAAGCTGTTATATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA 773  
DB 601 AAAGCTGTTATATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA 660  
QY 774 ATTTTAATATACCTAGTTATAGCATTTTCACTTTTGGATACCTTTTGGCAATTCATATC 833  
DB 661 ATTTTAATATACCTAGTTATAGCATTTTCACTTTTGGATACCTTTTGGCAATTCATATC 720  
QY 834 GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATGATACT 893  
DB 721 GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATGATACT 780  
QY 894 GCTTTTGGCTTCTCTGGTTTCTCTATATACAGTTTAAATTTTCTTATGTCCTTCTT 953  
DB 781 GCTTTTGGCTTCTCTGGTTTCTCTATATACAGTTTAAATTTTCTTATGTCCTTCTT 840  
QY 954 ATGGCAGTCAATGGCAGACTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTATTT 1013  
DB 841 ATGGCAGTCAATGGCAGACTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTATTT 900  
QY 1014 CTGCTTTTCTTTTCCNTTATGGAATATCATCTGTATTTTGTGCTTTTAAATGCTGACACTCTT 1073  
DB 901 CTGCTTTTCTTTTCCNTTATGGAATATCATCTGTATTTTGTGCTTTTAAATGCTGACACTCTT 960  
QY 1074 TTTAAAAATCAAAACATGTGGAAATAGTTGAAATTTTGTACTGTGGCTTTTGGATTT 1133  
DB 961 TTTAAAAATCAAAACATGTGGAAATAGTTGAAATTTTGTACTGTGGCTTTTGGATTT 1020  
QY 1134 ATTGSCCTTATGATAATCCTCATAGAAAGTTTCCCAAAATCGTTAGTGTGCTTTTCAGT 1193  
DB 1021 ATTGSCCTTATGATAATCCTCATAGAAAGTTTCCCAAAATCGTTAGTGTGCTTTTCAGT 1080  
QY 1194 CTTTCTGCTACTGTACTTTTGTGATTTGATTTGCAAGGTCATGCAATTTAGAGATTTT 1253  
DB 1081 CTTTCTGCTACTGTACTTTTGTGATTTGATTTGCAAGGTCATGCAATTTAGAGATTTT 1140  
QY 1254 AATGAAGTGTCTCATTTTCAAAATTTGACTGCAGGCCCATATCTCTAAATTTATACAAAT 1313  
DB 1141 AATGAAGTGTCTCATTTTCAAAATTTGACTGCAGGCCCATATCTCTAAATTTATACAAAT 1200  
QY 1314 ATCATGCTCACATTAATAGTATATCTATGTCCTCTTGGCTGTCTATCTTGATCAAGTC 1373

DB 1201 ATCATGCTCACATTAATAGTATATTTCTATGTCCTCTTGGCTGTCTATCTTGATCAAGTC 1260  
QY 1374 ATTCCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCAATTTGG 1433  
DB 1261 ATTCCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCAATTTGG 1320  
QY 1434 TCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGGCAATGTTAATGGAATATTTAGT 1493  
DB 1321 TCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGGCAATGTTAATGGAATATTTAGT 1380  
QY 1494 TTTAGTGAATTAATTTGAGCCAGTTTCTTCCAGAAATTTGTAGGAAAAGAACCCATAAGAAAT 1553  
DB 1381 TTTAGTGAATTAATTTGAGCCAGTTTCTTCCAGAAATTTGTAGGAAAAGAACCCATAAGAAAT 1440  
QY 1554 AGTGGTATTTCAAGAGCATATACAGAAAGAGGCTGAAATGTGGAGGCTTTGAGAAATTTG 1613  
DB 1441 AGTGGTATTTCAAGAGCATATACAGAAAGAGGCTGAAATGTGGAGGCTTTGAGAAATTTG 1500  
QY 1614 TCATTTGACATATATGAGGGTTCAGATTACTGCTTACTTTGGCCACAGTCGAAACAGGAAAG 1673  
DB 1501 TCATTTGACATATATGAGGGTTCAGATTACTGCTTACTTTGGCCACAGTCGAAACAGGAAAG 1560  
QY 1674 AGTACATTTGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1733  
DB 1561 AGTACATTTGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1620  
QY 1734 TATGGACACAGAGTCTTCAGAAATAGATGAATGTTTGAAGCAAGAAAATGATGGCAAT 1793  
DB 1621 TATGGACACAGAGTCTTCAGAAATAGATGAATGTTTGAAGCAAGAAAATGATGGCAAT 1680  
QY 1794 TGTCCACAGTTAGATATACACITTTGATGTTTGTAGAGTAGAAGAAAATTTATCAATTTG 1853  
DB 1681 TGTCCACAGTTAGATATACACITTTGATGTTTGTAGAGTAGAAGAAAATTTATCAATTTG 1740  
QY 1854 GCTTCAATCAAGAGGATACACAGCCAACTAATATACAGAGAGTCGAGAGGTTTACTTA 1913  
DB 1741 GCTTCAATCAAGAGGATACACAGCCAACTAATATACAGAGAGTCGAGAGGTTTACTTA 1800  
QY 1914 GATTTAGACATCCAGACTATCAAGATAACCAAGCTAAAAATTAAGTGGTGGTCAAAAA 1973  
DB 1801 GATTTAGACATCCAGACTATCAAGATAACCAAGCTAAAAATTAAGTGGTGGTCAAAAA 1860  
QY 1974 AGAAGCTGTCAATAGGAATGCTGCTTGTGGGACCCAAAGATCTGCTAGATGAA 2033  
DB 1861 AGAAGCTGTCAATAGGAATGCTGCTTGTGGGACCCAAAGATCTGCTAGATGAA 1920  
QY 2034 CCAACAGCTGGAAATGGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAAAATACAGA 2093  
DB 1921 CCAACAGCTGGAAATGGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAAAATACAGA 1980  
QY 2094 AAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTTGA 2153  
DB 1981 AAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTTGA 2040  
QY 2154 GATAGAAAGCTGTGATATCAAGAAATGCTGAAATGTTGTTGTTCTTCAATGTTCTCTC 2213  
DB 2041 GATAGAAAGCTGTGATATCAAGAAATGCTGAAATGTTGTTGTTCTTCAATGTTCTCTC 2100  
QY 2214 AAAAGTAAATGGGGATCGGCTACCGCTGAGCTGTACATAGACAAATTTTGTGCCACA 2273  
DB 2101 AAAAGTAAATGGGGATCGGCTACCGCTGAGCTGTACATAGACAAATTTTGTGCCACA 2160  
QY 2274 GAATCTCTTTCTTCTACTGTTAAACAAACATATACCTGGAGCTACTTTTATCAACAGAAAT 2333  
DB 2161 GAATCTCTTTCTTCTACTGTTAAACAAACATATACCTGGAGCTACTTTTATCAACAGAAAT 2220  
QY 2334 GACCAACAACTTGTGTATAGTGTGCTTTCAAGGACATGGACAAATTTTCAAGTTTGT 2393  
DB 2221 GACCAACAACTTGTGTATAGTGTGCTTTCAAGGACATGGACAAATTTTCAAGTTTGT 2280  
QY 2394 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCTTTCTTATGGTGTCTTCCATGACGACT 2453



Db 2281 TCTGCCCTAGACAGTCATCAAAATTTGGGTGGCAATTTCTTAGGGGTTTCATCAGCACT 2340  
Qy 2454 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTGACCAAGCAGATATAGT 2513  
Db 2341 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTGACCAAGCAGATATAGT 2400  
Qy 2514 GTATTTACTCAGCAGCCTGAGAGAAATGGAATCABAATCTTTTGATGAATGGA 2573  
Db 2401 GTATTTACTCAGCAGCCTGAGAGAAATGGAATCABAATCTTTTGATGAATGGA 2460  
Qy 2574 CAGAGCTTACTTATTTCTGAAACCAAGGCTCTCTAGTGAGCACCATGAGCCTTTGG 2633  
Db 2461 CAGAGCTTACTTATTTCTGAAACCAAGGCTCTCTAGTGAGCACCATGAGCCTTTGG 2520  
Qy 2634 AAACAACAGATGATACATAGCAAGTTTCATTTCTTTTACCTTGAAACGTGAAGTAAA 2693  
Db 2521 AAACAACAGATGATACATAGCAAGTTTCATTTCTTTTACCTTGAAACGTGAAGTAAA 2580  
Qy 2694 TCAGTGAGATCAGTGTGCTTCTGCTTTAAATTTTTTTCACAGTTCAGATTTTATGTTT 2753  
Db 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTAAATTTTTTTCACAGTTCAGATTTTATGTTT 2640  
Qy 2754 TTGGTTCACTCTTTTAAAAATGCTGTGGTTCCCATCAAACCTTTGTCAGACTATAT 2813  
Db 2641 TTGGTTCACTCTTTTAAAAATGCTGTGGTTCCCATCAAACCTTTGTCAGACTATAT 2700  
Qy 2814 TTCTTAAACCTGGAGACAAACCAATAAATACAAAACAAGTCTGCTTCTTCAAATTTCT 2873  
Db 2701 TTCTTAAACCTGGAGACAAACCAATAAATACAAAACAAGTCTGCTTCTTCAAATTTCT 2760  
Qy 2874 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAAGCCAGAACATATGTTGACG 2933  
Db 2761 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAAGCCAGAACATATGTTGACG 2820  
Qy 2934 ATGATTAATGACAGTACTATGATTCGCTGGCTCCCATAGTGGGCTTAAATGTGATG 2993  
Db 2821 ATGATTAATGACAGTACTATGATTCGCTGGCTCCCATAGTGGGCTTAAATGTGATG 2880  
Qy 2994 CATTCAGAAAGGACTATGTTTTTGCAGCTGTTTTCAACAGTACTATGTTTTTCAATTTA 3053  
Db 2881 CATTCAGAAAGGACTATGTTTTTGCAGCTGTTTTCAACAGTACTATGTTTTTCAATTTA 2940  
Qy 3054 CCTATATTAGTAATCATATAGTAATCTATCTTTTATCATTTAAATGTGACTGAAACC 3113  
Db 2941 CCTATATTAGTAATCATATAGTAATCTATCTTTTATCATTTAAATGTGACTGAAACC 3000  
Qy 3114 ATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTAAGTATAGTATGTTTTTAAATTTGAG 3173  
Db 3001 ATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTAAGTATAGTATGTTTTTAAATTTGAG 3060  
Qy 3174 CTGTATTTTCAAGCAGCTTTGCTTGGATCATGTTTACTGCAATGCCACCTTACTTTGCC 3233  
Db 3061 CTGTATTTTCAAGCAGCTTTGCTTGGATCATGTTTACTGCAATGCCACCTTACTTTGCC 3120  
Qy 3234 ATGGAATAATCGAGAAATCATAGATCAAGCTTATCTCAACTTAAACTTTTCAAGTCTT 3293  
Db 3121 ATGGAATAATCGAGAAATCATAGATCAAGCTTATCTCAACTTAAACTTTTCAAGTCTT 3180  
Qy 3294 TTGCCATCTGCATATTTGGATTTGGCAAGCTGTTGTTGATATCCCTTATTTTTTATCATTT 3353  
Db 3181 TTGCCATCTGCATATTTGGATTTGGCAAGCTGTTGTTGATATCCCTTATTTTTTATCATTT 3240  
Qy 3354 CTTATTTGATGCTAGAGCTTATGGCATTTTCAATTTGATTTATTTTATCTGTA 3413  
Db 3241 CTTATTTGATGCTAGAGCTTATGGCATTTTCAATTTGATTTATTTTATCTGTA 3300  
Qy 3414 AGTTCTCTGCTGTTTTTGGCTTATTTGTTATGTTTCCATCAGTATTTCTGTTCACT 3473  
Db 3301 AGTTCTCTGCTGTTTTTGGCTTATTTGTTATGTTTCCATCAGTATTTCTGTTCACT 3360  
Qy 3474 TATATTGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCATTTATC 3533  
Db 3361 TATATTGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCATTTATC 3420

Qy 3534 TATTTCTGTGGCAGCGTTGGCTTTGTATTGCAATCACTGAATAAATCTTTCTTTATGGATAC 3593  
Db 3421 TATTTCTGTGGCAGCGTTGGCTTTGTATTGCAATCACTGAATAAATCTTTCTTTATGGATAC 3480  
Qy 3594 ACAATTTGCAACTATTTCTTCAATTTATGCTTTTGTATCATCATTTCCAATCTATCCACTTTA 3653  
Db 3481 ACAATTTGCAACTATTTCTTCAATTTATGCTTTTGTATCATCATTTCCAATCTATCCACTTTA 3540  
Qy 3654 GGTGGCTCATTTCTTTTCAATAAAGATTTCTTGGAAAGATGTACGAAAAAATGTGACACCC 3713  
Db 3541 GGTGGCTCATTTCTTTCAATAAAGATTTCTTGGAAAGATGTACGAAAAAATGTGACACCC 3600  
Qy 3714 TATAATCCATGGATAGGCTTTTCAGTAGCTGTTATATCGCCTTACTCGCAGTGTGTACTG 3773  
Db 3601 TATAATCCATGGATAGGCTTTTCAGTAGCTGTTATATCGCCTTACTCGCAGTGTGTACTG 3660  
Qy 3774 TGGATTTTCTCTTCAATACTATGAGAAAAAATATGGAGGAGATCAATAAGAAAAAGAT 3833  
Db 3661 TGGATTTTCTCTTCAATACTATGAGAAAAAATATGGAGGAGATCAATAAGAAAAAGAT 3720  
Qy 3834 CCCTTTTTCAGAAACCTTTTCAACGAGTCTAAAAATAGGAAGCTTCCAGAACCAACAGAC 3893  
Db 3721 CCCTTTTTCAGAAACCTTTTCAACGAGTCTAAAAATAGGAAGCTTCCAGAACCAACAGAC 3780  
Qy 3894 AATGAGATCAAGATGTAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGGT 3953  
Db 3781 AATGAGATGTAAGATGTAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGGT 3840  
Qy 3954 TGCAGTGTGTGAGAGAGAACCATTCATATTTGGTCAGCAATTTGCATTAAGAAATATGAT 4013  
Db 3841 TGCAGTGTGTGAGAGAGAACCATTCATATTTGGTCAGCAATTTGCATTAAGAAATATGAT 3900  
Qy 4014 GACAAAGAAAGATTTTCTTCTTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 4073  
Db 3901 GACAAAGAAAGATTTTCTTCTTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 3960  
Qy 4074 TTCTGTGTGAAAAAGAGAGATCTTTAGGACTATTTGGGTCAAAATTTGGTCTGGCAAAAGC 4133  
Db 3961 TTCTGTGTGAAAAAGAGAGATCTTTAGGACTATTTGGGTCAAAATTTGGTCTGGCAAAAGC 4020  
Qy 4134 ACAATTTATTAATTTCTGTTGCTGATATTGAACCAACTTCAGGCCAGGTATTTTATGGA 4193  
Db 4021 ACAATTTATTAATTTCTGTTGCTGATATTGAACCAACTTCAGGCCAGGTATTTTATGGA 4080  
Qy 4194 GATTATTTCTCAGAGACAAAGTGAAAGATGATTTCACTGAAAGTGTATGGGTTACTGTCT 4253  
Db 4081 GATTATTTCTCAGAGACAAAGTGAAAGTGTATTTCACTGAAAGTGTATGGGTTACTGTCT 4140  
Qy 4254 CAGATAAACCCCTTTGTCGCGCAGATCTCATTTGCGAGAACATTTTGAATTTATGAGCT 4313  
Db 4141 CAGATAAACCCCTTTGTCGCGCAGATCTCATTTGCGAGAACATTTTGAATTTATGAGCT 4200  
Qy 4314 GTCAAGGAATAGTCAAGTGACATGAAGAGTCAATAGTCGAATAACATGACATTT 4373  
Db 4201 GTCAAGGAATAGTCAAGTGACATGAAGAGTCAATAGTCGAATAACATGACATTT 4260  
Qy 4374 GATTTTAAAGAACATCTTCAGAGAGCTGTAAAGAAATACCTCGCAGGAATCAAAACGAAAG 4433  
Db 4261 GATTTTAAAGAACATCTTCAGAGAGCTGTAAAGAAATACCTCGCAGGAATCAAAACGAAAG 4320  
Qy 4434 TTGTGTTTGTCTCTAAGTATGCTTAGGAAATCTTCAGATTTACTTTTGTAGTAAGAACATCT 4493  
Db 4321 TTGTGTTTGTCTCTAAGTATGCTTAGGAAATCTTCAGATTTACTTTTGTAGTAAGAACATCT 4380  
Qy 4494 ACAGGTATGATCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAATTCGATTTAAA 4553  
Db 4381 ACAGGTATGATCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAATTCGATTTAAA 4440  
Qy 4554 AACAGAAAGCGGCTGCTATTTCTGACCCTCACTATATGAGAGAGCAGAGGCTGTCTGT 4613  
Db 4441 AACAGAAAGCGGCTGCTATTTCTGACCCTCACTATATGAGAGAGCAGAGGCTGTCTGT 4500

QY 4614 GATCAGTAGCTATCATGCTCTCTGGCAGTAAAGATGATCGAAGCAGTACAACTCTA 4673  
 Db |||||  
 QY 4501 GATCAGTAGCTATCATGCTCTCTGGCAGTAAAGATGATCGAAGCAGTACAACTCTA 4560  
 Db |||||  
 QY 4674 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGGACTGATAGAAAAC 4733  
 Db |||||  
 QY 4561 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGGACTGATAGAAAAC 4620  
 Db |||||  
 QY 4734 CTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 4793  
 Db |||||  
 QY 4621 CTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 4680  
 Db |||||  
 QY 4794 GAAAGTCTTTCTCTATTTTGGCTTATAAAATTCCTAAGGAGATGTTCACTCCCTTCA 4853  
 Db |||||  
 QY 4681 GAAAGTCTTTCTCTATTTTGGCTTATAAAATTCCTAAGGAGATGTTCACTCCCTTCA 4740  
 Db |||||  
 QY 4854 CAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTCCCATTTGAAGATATAGCTTT 4913  
 Db |||||  
 QY 4741 CAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTCCCATTTGAAGATATAGCTTT 4800  
 Db |||||  
 QY 4914 TCTCAAGCAACATTTGGAAGAGCTTTTGTAGAACTCACTAAAGCAAGAGGAGGAGAT 4973  
 Db |||||  
 QY 4801 TCTCAAGCAACATTTGGAAGAGCTTTTGTAGAACTCACTAAAGCAAGAGGAGGAGAT 4860  
 Db |||||  
 QY 4974 AATAGTTGTGAAGCTTTAAACAGACACACTTTTGTGGGAACGAAACACAAAGAGATAGTA 5033  
 Db |||||  
 QY 4861 AATAGTTGTGAAGCTTTAAACAGACACACTTTTGTGGGAACGAAACACAAAGAGATAGTA 4920  
 Db |||||  
 QY 5034 GTATTTGA 5042  
 Db |||||  
 QY 4921 GTATTTGA 4929  
 Db |||||

RESULT 7  
 ABSS7750  
 ID ABSS7750 standard; cDNA; 4917 BP.  
 XX  
 AC ABSS7750;  
 XX  
 DT 04-FEB-2003 (first entry)  
 XX  
 DE Coding sequence of human ATP binding cassette ABCA5 transporter #1.  
 XX  
 KW Human; ATP binding cassette; ABC; ABCA5; transporter;  
 KW neurotoxin transport; beta-amyloid peptide; chromosome mapping;  
 KW blood brain barrier transport; tissue typing; predictive medicine;  
 KW ABCA5 mediated disorder; ABCA5 related disorder; gene therapy; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..4917  
 FT /\*tag= a  
 FT /product= "ABCA5"  
 FT /note= "ATP binding cassette (ABC) A5 transporter"  
 FT /transl\_except= (pos:3160..3162, aa:Ser)  
 XX  
 FN US2002123107-A1.  
 XX  
 PD 05-SEP-2002.  
 XX  
 PF 01-MAR-2002; 2002US-00090458.  
 XX  
 PR 02-MAR-2001; 2001US-0272885P.  
 XX  
 PA (ACT1-) ACTIVE PASS PHARM INC.  
 XX  
 PI Chen H, Kilinski L, Le Bihan S;  
 XX  
 XX WPI; 2003-065798/06.  
 DR P-PSDB; ABG72423.  
 XX  
 PT Novel isolated ATP binding cassette transporter family polypeptide,

PT ABCA5, useful for treating disorders associated with aberrant or unwanted  
 PT ABCA5 transporter expression or activity.  
 XX  
 PS Claim 2; Fig 1; 52pp; English.  
 XX  
 CC The invention describes an isolated ATP binding cassette (ABC)  
 CC transporter family polypeptide (I), designated ABCA5. (I) or the  
 CC polynucleotide encoding it (II) are useful as targets for developing  
 CC modulating agents to regulate a variety of cellular processes,  
 CC particularly the transport of neurotoxic molecules, e.g., beta-amyloid  
 CC peptide (Abeta), across cell membranes or, e.g., the blood brain barrier  
 CC (BBB), as targets for developing modulating agents of multi-drug  
 CC resistance as diagnostic and therapeutic tools, or to treat disorders  
 CC associated with aberrant or unwanted ABCA5 transporter expression or  
 CC activity. (I), (II) or a host cell (III) expressing (II) are useful in  
 CC screening assays, detection assays (e.g., chromosomal mapping, tissue  
 CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,  
 CC prognostic assays, monitoring clinical trials and pharmacogenomics), and  
 CC in methods of treatment (e.g., therapeutic and prophylactic). (I) or  
 CC (III) are useful as reagents or targets in assays applicable to treatment  
 CC and diagnosis of ABCA5-mediated or related disorders. (I) is useful to  
 CC screen for naturally occurring ABCA5 substrates; to screen for drugs or  
 CC compounds which modulate ABCA5 activity; as a bait protein in a yeast two  
 CC -hybrid or three-hybrid assay; and to identify other proteins which bind  
 CC to or interact with ABCA5. (II) is useful in: gene therapy; to detect  
 CC ABCA5 mRNA or a genetic alteration in a ABCA5 gene; to modulate ABCA5  
 CC activity; to locate gene regions associated with genetic disease or to  
 CC associate ABCA5 with the disease, to identify an individual from a minute  
 CC biological sample (tissue typing), and to aid in forensic identification  
 CC of the biological sample. This sequence encodes a novel human ATP binding  
 CC cassette (ABC) A5 transporter  
 XX  
 SQ Sequence 4917 BP; 1540 A; 821 C; 955 G; 1599 T; 0 U; 2 Other;

Query Match 87.9%; Score 4811.6; DB 7; Length 4917;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 4894; Conservative 0; Mismatches 4; Indels 50; Gaps 3;  
 QY 114 ATGTCCTACGCAATTTAGGAGGTAGGAGCTTTGGAGACAGACAGACCACTTCTACTGAAG 173  
 Db |||||  
 QY 174 AATTAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 233  
 Db |||||  
 QY 61 AATTAATTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 120  
 Db |||||  
 QY 234 TTTTCTTTTATTTTGGTTAATTAATTAATGAGTAGTGTTCAGGAAATCTTTTCCACTA 293  
 Db |||||  
 QY 121 TTTTCTTTTATTTTGGTTAATTAATTAATGAGTAGTGTTCAGGAAATCTTTTCCACTA 180  
 Db |||||  
 QY 294 GTGCCTAATATAGAACTCAATCCTTATGGAACAAGTTTACTCTTTCTAATCTAATCTTGGGA 353  
 Db |||||  
 QY 181 GTGCCTAATATAGAACTCAATCCTTATGGAACAAGTTTACTCTTTCTAATCTAATCTTGGGA 240  
 Db |||||  
 QY 354 TATATCTCAGTGACTAATATTACAGCAGCATCATGCAAGAGTGTCTACTGATCATCTA 413  
 Db |||||  
 QY 241 TATATCTCAGTGACTAATATTACAGCAGCATCATGCAAGAGTGTCTACTGATCATCTA 300  
 Db |||||  
 QY 414 COTGATGTCTAATATTACTGAAGAATATACAAATGAAAAAGAAATGTTTAACTCCAGTCTC 473  
 Db |||||  
 QY 301 COTGATGTCTAATATTACTGAAGAATATACAAATGAAAAAGAAATGTTTAACTCCAGTCTC 360  
 Db |||||  
 QY 474 TCTTAGCCGAGCACTTTGTAGGTGTGTTTCAAGACTCCATGCTCTCTATGAACTTCGT 533  
 Db |||||  
 QY 361 TCTTAGCCGAGCACTTTGTAGGTGTGTTTCAAGACTCCATGCTCTCTATGAACTTCGT 420  
 Db |||||  
 QY 534 TTTTCTCTGTATGATTCAGTATCTTCTATTTATATGAGATTCAAGAGCTGGCTGTTC 593  
 Db |||||  
 QY 421 TTTTCTCTGTATGATTCAGTATCTTCTATTTATATGAGATTCAAGAGCTGGCTGTTC 480  
 Db |||||  
 QY 594 AATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTCAGGTTTACAGATCCATA 653  
 Db |||||  
 QY 481 AATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTACAGATCCATA 540  
 Db |||||

QY 654 GATGCTGCCATATACAGTTGAAGCAACATGTTTCTCTTTGGAGGAGCTGGAGTCAACT 713  
DB 541 GATGCTGCCATATACAGTTGAAGCAACATGTTTCTCTTTGGAGGAGCTGGAGTCAACT 600  
QY 714 AAAGCTGTATATATGGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGATA 773  
DB 601 AAAGCTGTATATATGGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGATA 660  
QY 774 ATTTTAATATACCTAGTTATAGCAATTTTCACTTTTGGATACCTTTTGGCAATTCATATC 833  
DB 661 ATTTTAATATACCTAGTTATAGCAATTTTCACTTTTGGATACCTTTTGGCAATTCATATC 720  
QY 834 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCAGATACCT 893  
DB 721 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCAGATACCT 780  
QY 894 GCCTTTTGGCTTCCCTGGTTCCTCTATATACAGTTTAAATTTTCTATGTCCTTCTT 953  
DB 781 GCCTTTTGGCTTCCCTGGTTCCTCTATATACAGTTTAAATTTTCTATGTCCTTCTT 840  
QY 954 ATGGCAGTCATTTGGCAGACCTTCTTCTATATTTCCCTCAAGTAGCAGCAATTTGATATTT 1013  
DB 841 ATGGCAGTCATTTGGCAGACCTTCTTCTATATTTTCCCTCAAGTAGCAGCAATTTGATATTT 900  
QY 1014 CTGCTTTTTCCTTTATGGAATATCACTGTATTTTGTCTTTTAAATGTCGACCTCTT 1073  
DB 901 CTGCTTTTTCCTTTATGGAATATCACTGTATTTTGTCTTTTAAATGTCGACCTCTT 960  
QY 1074 TTTAAAAAATCAAAACATGTGGGAATAGTTGAAATTTTGTACTGTGGCTTTTGGATTT 1133  
DB 961 TTTAAAAAATCAAAACATGTGGGAATAGTTGAAATTTTGTACTGTGGCTTTTGGATTT 1020  
QY 1134 ATGGCTTATGATAATCTCATAGAAAGTTTCCCAAAATCGTGTAGTGGCTTTTCAAT 1193  
DB 1021 ATGGCTTATGATAATCTCATAGAAAGTTTCCCAAAATCGTGTAGTGGCTTTTCAAT 1080  
QY 1194 CTTTCTGTCTACTGCTTTTGTGATTTGATTTGCAAGCTCATGCAATTTAGCAATTT 1253  
DB 1081 CTTTCTGTCTACTGCTTTTGTGATTTGATTTGCAAGCTCATGCAATTTAGCAATTT 1140  
QY 1254 AATGAAGTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCTCTTAATTTTCAAT 1313  
DB 1141 AATGAAGTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCTCTTAATTTTCAAT 1200  
QY 1314 ATCATGCTCACACTTAATAGTATATCTATGCTCTCTGCTCTCTATCTGATCAAGTC 1373  
DB 1201 ATCATGCTCACACTTAATAGTATATCTATGCTCTCTGCTCTCTATCTGATCAAGTC 1260  
QY 1374 ATTCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTTTCAATTTGG 1433  
DB 1261 ATTCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTTTCAATTTGG 1320  
QY 1434 TCRAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCAATGTTTAAATGGAATATTTAGT 1493  
DB 1321 TCRAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCAATGTTTAAATGGAATATTTAGT 1380  
QY 1494 TTTAGTGAATTTATGAGGAGTTTCTTCAAGATTTTGTAGGAAAAAGAGCCATAAGATTT 1553  
DB 1381 TTTAGTGAATTTATGAGGAGTTTCTTCAAGATTTTGTAGGAAAAAGAGCCATAAGATTT 1440  
QY 1554 AGTGGTATTCAGAGACATACAGAAAGAGGGTGAATAATGTCGAGCTTTTGGAAATTTG 1613  
DB 1441 AGTGGTATTCAGAGACATACAGAAAGAGGGTGAATAATGTCGAGCTTTTGGAAATTTG 1500  
QY 1614 TCATTTGACATATATGAGGGTTCAGATTTCTGCTTACTTTGGCCACAGTGGGAACAGGAAAG 1673  
DB 1501 TCATTTGACATATATGAGGGTTCAGATTTCTGCTTACTTTGGCCACAGTGGGAACAGGAAAG 1560  
QY 1674 AGTACATTTGATGAATTTCTTTGCTGAGCTCTGCCACCTTCTGATGGGTTTGGCATCTATA 1733  
DB 1561 AGTACATTTGATGAATTTCTTTGCTGAGCTCTGCCACCTTCTGATGGGTTTGGCATCTATA 1620

QY 1734 TATGGACACAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTTGGCATTT 1793  
DB 1621 TATGGACACAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTTGGCATTT 1680  
QY 1794 TGTCCACAGTTAGATATACACTTTGATGTTTGAAGTGTAGAGAGAAAAATTTTCAATTTTGG 1853  
DB 1681 TGTCCACAGTTAGATATACACTTTGATGTTTGAAGTGTAGAGAGAAAAATTTTCAATTTTGG 1740  
QY 1854 GCTTCAATCAAGAGGATACCCAGCCAAATAATAAATACAGAAAGTGCAGAAAGGTTTACTTA 1913  
DB 1741 GCTTCAATCAAGAGGATACCCAGCCAAATAATAAATACAGAAAGTGCAGAAAGGTTTACTTA 1800  
QY 1914 GATTTAGACATGACAGACTATCAAGATACCAAGCTTAAAAATTAAGTGGTGGTCAAAAA 1973  
DB 1801 GATTTAGACATGACAGACTATCAAGATACCAAGCTTAAAAATTAAGTGGTGGTCAAAAA 1860  
QY 1974 AGAAAGCTGTCAATTAGGAATTTGCTGTTTGGGAACCCAAAGATACCTGCTGTAGATGAA 2033  
DB 1861 AGAAAGCTGTCAATTAGGAATTTGCTGTTTGGGAACCCAAAGATACCTGCTGTAGATGAA 1920  
QY 2034 CCACAGCTGGATGGAATGGAACCTGTTCTGACATATTTGTATGGAATCTTTTAAAAATACAGA 2093  
DB 1921 CCACAGCTGGATGGAATGGAACCTGTTCTGACATATTTGTATGGAATCTTTTAAAAATACAGA 1980  
QY 2094 AAAGCAATCGGGTGACAGTGTTCAGTACTCTATTTCTATGATGAAGCTGACATTTCTTGA 2153  
DB 1981 AAAGCAATCGGGTGACAGTGTTCAGTACTCTATTTCTATGATGAAGCTGACATTTCTTGA 2040  
QY 2154 GATAGGAAGCTGTGATATCAAGGAATGCTGGAATGTTGTTTCAATGTTCTTCAATGTTCTC 2213  
DB 2041 GATAGGAAGCTGTGATATCAAGGAATGCTGGAATGTTGTTTCAATGTTCTTCAATGTTCTC 2100  
QY 2214 AAAGTAATCGGGGATCGCTACCGCTTGACATGTACATAGACAAAAATTTTGTGCGACA 2273  
DB 2101 AAAGTAATCGGGGATCGCTACCGCTTGACATGTACATAGACAAAAATTTTGTGCGACA 2160  
QY 2274 GAATCTCTTCTTCACTGTTTAAACAACATATACCTGGAGCTTACTTTATTAACAAGAT 2333  
DB 2161 GAATCTCTTCTTCACTGTTTAAACAACATATACCTGGAGCTACTTTATTAACAAGAT 2220  
QY 2334 GACCAACAACTTTGTATAGCTTGCCTTTCAAGGACATGGACAAATTTTCAAGTTTGT 2393  
DB 2221 GACCAACAACTTTGTATAGCTTGCCTTTCAAGGACATGGACAAATTTTCAAGTTTGT 2280  
QY 2394 TCTGCCCCATAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGTTTCCATGACGACT 2453  
DB 2281 TCTGCCCCATAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGTTTCCATGACGACT 2340  
QY 2454 TTGGAAGACGTAATTTTAAAGCTAGAGTTGAAGCAGAAAAATGACCAAGACAGATTTAGT 2513  
DB 2341 TTGGAAGACGTAATTTTAAAGCTAGAGTTGAAGCAGAAAAATGACCAAGACAGATTTAGT 2400  
QY 2514 GTATTTTACTCAGCAGCCACTGGAGGAGAAATGGAATTCAAAATCTTTTGTAGTGAATGGA 2573  
DB 2401 GTATTTTACTCAGCAGCCACTGGAGGAGAAATGGAATTCAAAATCTTTTGTAGTGAATGGA 2460  
QY 2574 CAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAGGACCAATGAGGCTTTGG 2633  
DB 2461 CAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAGGACCAATGAGGCTTTGG 2520  
QY 2634 AAACAACAGATCTATACAAATAGCAAGTTTCAATTTCTTACCTTTGAAACGTTGAAGTAAA 2693  
DB 2521 AAACAACAGATCTATACAAATAGCAAGTTTCAATTTCTTACCTTTGAAACGTTGAAGTAAA 2580  
QY 2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2753  
DB 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2640  
QY 2754 TTGTTTCACTCTTTTAAAAATGCTGCTGTTTCCATCAAACTTGTTCAGAGCTTATAT 2813  
DB 2641 TTGTTTCACTCTTTTAAAAATGCTGCTGTTTCCATCAAACTTGTTCAGAGCTTATAT 2700  
QY 2814 TTTCTTAAACCTGGAGACAAACCAACATATAACAAACAGTCTGCTTCTTCAAAATTTCT 2873

Db 2701 TTTCTAAACCTGGAGACAAACCCATATAAATACAAAACAAGTCGTCTTCTTCAAAATCT 2760  
Qy 2874 GCT-----GACTCAGATATCAGTGATCTTAATAGCTTTTTCACAAGC 2915  
Db 2761 GCTGGTGAGAGTGNNNGAAGACTCAGATATCAGTGATCTTAATAGCTTTTTCACAAGC 2820  
Qy 2916 CAGAACATAATGGTGACGATGATTAATGACAGTGACTATGATCCGTCGCTCCCAATAGT 2975  
Db 2821 CAGAACATAATGGTGACGATGATTAATGACAGTGACTATGATCCGTCGCTCCCAATAGT 2880  
Qy 2976 CGCGCTTTAAATGTGATGCAATCAGAAAAGACTATGTTTTTGCAGCTGTTTTCAACAGT 3035  
Db 2881 CGCGCTTTAAATGTGATGCAATCAGAAAAGACTATGTTTTTGCAGCTGTTTTCAACAGT 2940  
Qy 3036 ACTATGGTTTATCTTTACCTATATATAGTGAATATCATTAGTAACACTACTATCTTTATCAT 3095  
Db 2941 ACTATGGTTTATCTTTACCTATATATAGTGAATATCATTAGTAACACTACTATCTTTATCAT 3000  
Qy 3096 TTAAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTAATCTGAT 3155  
Db 3001 TTAATGTGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTAATCTGAT 3060  
Qy 3156 ATAGTTTTTAAATGTGACTGATTTTCAAGCAGCTTTGCTGG-ATCATGTGTTACTGC 3214  
Db 3061 ATAGTTTTTAAATGTGACTGATTTTCAAGCAGCTTTGCTTTGGAATCATTTGTAACCTGC 3120  
Qy 3215 AATGCCACTTACTTTGCCATGGAAATGCAGAGATCATAGATCAAAAGCTTATACTCA 3274  
Db 3121 AATGCCACTTACTTTGCCATGGAAATGCAGAGATCAT----- 3161  
Qy 3275 ACTTAAACCTTTCAAGTCTTTTGGCCATCTGCATATTTGGATTTGGACAAGCTGTTGTGATAT 3334  
Db 3162 -----AGGCTTTTGGCCATCTGCATATTTGGATTTGGACAAGCTGTTGTGATAT 3209  
Qy 3335 CCCCTATTTTATCATCTTATTTTTCATGCTAGGAGCTTATTTGGCAATTCATTTATGG 3394  
Db 3210 CCCCTATTTTATCATCTTATTTTTCATGCTAGGAGCTTATTTGGCAATTCATTTATGG 3269  
Qy 3395 ATTATATTTTATCTGTAAGTTCCTTTGCTGTGTTTTTTTGGCCTTATGTTTATGTTCC 3454  
Db 3270 ATTATATTTTATCTGTAAGTTCCTTTGCTGTGTTTTTTTGGCCTTATGTTTATGTTCC 3329  
Qy 3455 ATCAGTATTTCTGTTCACTTATTTGCTTTTTCACCTTTAAGAAATTTTAAATACCAA 3514  
Db 3330 ATCAGTATTTCTGTTCACTTATTTGCTTTTTCACCTTTAAGAAATTTTAAATACCAA 3389  
Qy 3515 AGAATTTTGGTCAATTTATCTATTTCTGTGGCAGCTTGGCTTGTATTGCAATCACTGAAT 3574  
Db 3390 AGAATTTTGGTCAATTTATCTATTTCTGTGGCAGCTTGGCTTGTATTGCAATCACTGAAT 3449  
Qy 3575 AACTTTCTTTATGGGATACAAATTTGCAACTATTTCTGATTTGCAATCACTGAAT 3634  
Db 3450 AACTTTCTTTATGGGATACAAATTTGCAACTATTTCTGATTTGCAATCACTGAAT 3509  
Qy 3635 TCCATCTATCCACTTCTAGTGTGCTGATTTCTTTCATAAGATTTCTTGGAAATGT 3694  
Db 3510 TCCATCTATCCACTTCTAGTGTGCTGATTTCTTTCATAAGATTTCTTGGAAATGT 3569  
Qy 3695 ACGAAAAATGTGGACACCTATAATCCATGGGATAGGCTTTTCAAGTGTATTATCGCC 3754  
Db 3570 ACGAAAAATGTGGACACCTATAATCCATGGGATAGGCTTTTCAAGTGTATTATCGCC 3629  
Qy 3755 TTACCTGCGATGTGTTACTGTGATTTTCTCTTACAAATCTATGAGAAATATGGAGG 3814  
Db 3630 TTACCTGCGATGTGTTACTGTGATTTTCTCTTACAAATCTATGAGAAATATGGAGG 3689  
Qy 3815 CAGATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAAGTGTCTTAAATATGGAA 3874  
Db 3690 CAGATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAAGTGTCTTAAATATGGAA 3749  
Qy 3875 GCTTCCAGAACCAACCTGAGACATATGAGATGAAGATGCAAGCTGAAGACTTAA 3934

Db 3750 GCTTCCAGAACCCACAGACAATAGGATGAAGATGTCAAAGCTGAAGACTAAA 3809  
Qy 3935 GGTCAAAAGAGCTGATGGGTGGCAGTGTGTGAGAGAAACCATCATTTATGTTGAGCAA 3994  
Db 3810 GGTCAAAAGAGCTGATGGGTGGCAGTGTGTGAGAGAAACCATCATTTATGTTGAGCAA 3869  
Qy 3995 TTTTGCATATAAGAAATATGATGACAAAGAAAGATTTTCTTCTTTCAAGAAAGTAAAGAAAGT 4054  
Db 3870 TTTTGCATATAAGAAATATGATGACAAAGAAAGATTTTCTTCTTTCAAGAAAGTAAAGAAAGT 3929  
Qy 4055 GGCACATAAATACATCTCTTTCTGTGTGAAAAGAGAGATCTTTAGACATTTGGGTCC 4114  
Db 3930 GGCACATAAATACATCTCTTTCTGTGTGAAAAGAGAGATCTTTAGACATTTGGGTCC 3989  
Qy 4115 AAATGCTGTGCAAAAAGCACAAATTAATAATTTCTGTTGGTGATATTGAACCAACTTC 4174  
Db 3990 AAATGCTGTGCAAAAAGCACAAATTAATAATTTCTGTTGGTGATATTGAACCAACTTC 4049  
Qy 4175 AGCCAGGTATTTTATAGAGATTTTCTTCAGACAGCAAGTGAAGATGATGATCTACTGAA 4234  
Db 4050 AGCCAGGTATTTTATAGAGATTTTCTTCAGACAGCAAGTGAAGATGATGATCTACTGAA 4109  
Qy 4235 GTGTATGGGTACTGTCTCAGATAAAACCTTTGTGGCCAGATATCTATATTGCGAGAAACA 4294  
Db 4110 GTGTATGGGTACTGTCTCAGATAAAACCTTTGTGGCCAGATATCTATATTGCGAGAAACA 4169  
Qy 4295 TTTTGAATTTTATGGAGCTGTCAAAGGAATGAGTGCAAGTGACATGAAGAAAGTCTATAAG 4354  
Db 4170 TTTTGAATTTTATGGAGCTGTCAAAGGAATGAGTGCAAGTGACATGAAGAAAGTCTATAAG 4229  
Qy 4355 TCGAATAACACATGCACTTGATTTTAAAGAAACATCTTCAGAAAGACTGTAAAGAAACTACC 4414  
Db 4230 TCGAATAACACATGCACTTGATTTTAAAGAAACATCTTCAGAAAGACTGTAAAGAAACTACC 4289  
Qy 4415 TGCAAGAAATCAAAACGAAAGTTGTTTTCTCTTAAGATGTGTTAGGAAATCTCAGATTAC 4474  
Db 4290 TGCAAGAAATCAAAACGAAAGTTGTTTTCTCTTAAGTATGTTAGGAAATCTCAGATTAC 4349  
Qy 4475 TTTTGTAGATGAACCATCTACAGGTATGATCCCAAGCCAAACACACATGTGGCGAGC 4534  
Db 4350 TTTTGTAGATGAACCATCTACAGGTATGATCCCAAGCCAAACACACATGTGGCGAGC 4409  
Qy 4535 AATTGCAACTGCAATTTTAAACACAGAAAGCGGCTGCTATTCTGACCACTCATATATGGA 4594  
Db 4410 AATTGCAACTGCAATTTTAAACACAGAAAGCGGCTGCTATTCTGACCACTCATATATGGA 4469  
Qy 4595 GGAGGACAGAGCTGTCTGTGATCGAGTATCATGTTCTGTGGCAGTTAAGATGAT 4654  
Db 4470 GGAGGACAGAGCTGTCTGTGATCGAGTATCATGTTCTGTGGCAGTTAAGATGAT 4529  
Qy 4655 CGAAACAGTACAAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTTAAAT 4714  
Db 4530 CGAAACAGTACAAACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTTAAAT 4589  
Qy 4715 GAAGGACTGGATAGAAAACCTAGAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTT 4774  
Db 4590 GAAGGACTGGATAGAAAACCTAGAGTAGACCGCTTCAAAGAGAGAAATTCAGTATATTTT 4649  
Qy 4775 CCCAATGTCAAGCGCTCAGGAAAGTTTCTCTATTTTGGCTTATAAAATTCCTAAGGA 4834  
Db 4650 CCCAATGTCAAGCGCTCAGGAAAGTTTCTCTATTTTGGCTTATAAAATTCCTAAGGA 4709  
Qy 4835 AGATGTTTCAGTCCCTTTTCAATCTTTTAAAGTGAAGAGCTAAACATGCTTTTGC 4894  
Db 4710 AGATGTTTCAGTCCCTTTTCAATCTTTTAAAGTGAAGAGCTAAACATGCTTTTGC 4769  
Qy 4895 CATTGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAA 4954  
Db 4770 CATTGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAA 4829  
Qy 4955 AGAACAGAGAGGAGAGATTAATAGTTGTGAACTTTTAAACAGCACTTTTGTGGGAACG 5014  
Db 4830 AGAACAGAGAGGAGAGATTAATAGTTGTGAACTTTTAAACAGCACTTTTGTGGGAACG 4889

QY 5015 AACACAGAGATAGTAGTATTTTGA 5042  
 Db 4890 AACACAGAGATAGTAGTATTTTGA 4917

RESULT 8  
 AAD37619  
 ID AAD37619 standard; cDNA; 4785 BP.  
 AC AAD37619;  
 XX AAD37619;  
 DT 10-SEP-2002 (first entry)  
 XX Human transporter protein cDNA #2.  
 DE Human; novel human protein; NHP; transporter protein; mental disorder;  
 KW cancer; gene therapy; drug screening; nutraceutical application; gene;  
 KW cosmetic application; polymorphism; ss.  
 XX Homo sapiens.  
 OS

XX Key Location/Qualifiers  
 FH 1. .4785  
 FT /tag= a  
 FT /product= "Human transporter protein #2"  
 FT /transl\_except= (pos:808. .810, aa:Arg)  
 FT /transl\_except= (pos:2494. .2496, aa:Ser)  
 FT /transl\_except= (pos:2878. .2880, aa:Val)  
 FT misc\_feature 810  
 FT /tag= b  
 FT /note= "This degenerate base represents a polymorphic site"  
 FT misc\_feature 2494  
 FT /tag= c  
 FT /note= "This degenerate base represents a polymorphic site"  
 FT misc\_feature 2878  
 FT /tag= d  
 FT /note= "This degenerate base represents a polymorphic site"  
 XX W0200231147-A2.  
 XX 18-APR-2002.  
 XX 04-OCT-2001; 2001WO-US031113.  
 XX 10-OCT-2000; 2000US-0239629P.  
 XX (LEXI-) LEXICON GENETICS INC.  
 XX Hu Y, Nepomnichy B;  
 XX WPI; 2002-454552/48.  
 XX P-PSDB; AAE23657.  
 XX Novel nucleic acid molecule encoding novel human proteins, useful for  
 PT therapeutic, diagnostic and pharmacogenomic applications.  
 XX Claim 1; Page 39-41; 46pp; English.  
 XX The present sequence is a cDNA coding for novel human protein (NHP),  
 CC human transporter protein. NHPs shares structural similarity with the  
 CC mammalian ATP-binding cassette (ABC) transporters and multidrug  
 CC resistance transporters. NHP polynucleotides are useful for the  
 CC therapeutic, diagnostic and pharmacogenomic applications. They are used  
 CC for detecting and treating mental disorders and cancers. They are also  
 CC used in gene therapy. NHP polypeptides are useful for diagnosis, drug  
 CC screening, clinical trial monitoring, treatment of diseases and  
 CC disorders, and cosmetic or nutraceutical applications  
 XX Sequence 4785 BP; 1489 A; 807 C; 917 G; 1568 T; 0 U; 4 Other;

Query Match 86.9%; Score 4759; DB 6; Length 4785;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 4759; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 114 ATGTCACCTGCAATAGGAGGTAGGAGTGGAGACAGACAGACACTTCTACTGGAAG 173  
 Db 1 ATGTCACCTGCAATAGGAGGTAGGAGTGGAGACAGACAGACACTTCTACTGGAAG 60

QY 174 AATTACTTAATTAATGAGAACCAAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 233  
 Db 61 AATTACTTAATTAATGAGAACCAAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 120

QY 234 TTTTCTTTTATTTTGGTTTATTAATAGCATGATGCATCCAAATAGAAATATGAAGA 293  
 Db 121 TTTTCTTTTATTTTGGTTTATTAATAGCATGATGCATCCAAATAGAAATATGAAGA 180

QY 294 GTGCCTAATATAGAACTCAATCCTTATGAGCAAGTTTACTCTTTCTAATTAATTTTGGGA 353  
 Db 181 GTGCCTAATATAGAACTCAATCCTTATGAGCAAGTTTACTCTTTCTAATTAATTTTGGGA 240

QY 354 TATACTCCAGTGACTAATATTACAGAGCATCATGAGAAAGTGTCTACTGATCATCTA 413  
 Db 241 TATACTCCAGTGACTAATATTACAGAGCATCATGAGAAAGTGTCTACTGATCATCTA 300

QY 414 CCGATGTCATTAATTAATGAGAAATATACAAATGAAAAGAAATGTTTAAACATCCAGTCTC 473  
 Db 301 CCGATGTCATTAATTAATGAGAAATATACAAATGAAAAGAAATGTTTAAACATCCAGTCTC 360

QY 474 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAGACTCCATGTCTCTATGAATTCGT 533  
 Db 361 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAGACTCCATGTCTCTATGAATTCGT 420

QY 534 TTTTCTCTGATATGATTCAGATCTCTTATTTATATGAAATCAAGAGTGGCTGTCTCA 593  
 Db 421 TTTTCTCTGATATGATTCAGATCTCTTATTTATATGAAATCAAGAGTGGCTGTCTCA 480

QY 594 AAATCATGTGAGGCTGCTCAGTACTGCTGCTCAGTTTCAAGTTTCAAGCATCCATA 653  
 Db 481 AAATCATGTGAGGCTGCTCAGTACTGCTGCTCAGTTTCAAGTTTCAAGCATCCATA 540

QY 654 GATGCTGCATATACAGTTGAAGACCAATGTTTCTTTTGGAGGAGCTGGAGTCAACT 713  
 Db 541 GATGCTGCATATACAGTTGAAGACCAATGTTTCTTTTGGAGGAGCTGGAGTCAACT 600

QY 714 AAAGCTGTTTATGAGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGTA 773  
 Db 601 AAAGCTGTTTATGAGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGTA 660

QY 774 ATTTAATATACCTAGTTATAGCAATTTTACCTTTTGGATACCTTTTGGCAATTCATATC 833  
 Db 661 ATTTAATATACCTAGTTATAGCAATTTTACCTTTTGGATACCTTTTGGCAATTCATATC 720

QY 834 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 893  
 Db 721 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 780

QY 894 GCCTTTTGGCTTCTCTGGGTTCTTCTATATACAGTTTAAATTTTCTATGTGCTCTCT 953  
 Db 781 GCCTTTTGGCTTCTCTGGGTTCTTCTATATACAGTTTAAATTTTCTATGTGCTCTCT 840

QY 954 ATGGCAGTCATTGGCAGAGCTTCTTGTATTTTCTCCTCAAAAGTAGCAGCATTTGATATTT 1013  
 Db 841 ATGGCAGTCATTGGCAGAGCTTCTTGTATTTTCTCCTCAAAAGTAGCAGCATTTGATATTT 900

QY 1014 CTGCTTTTCTTCTTATGAGTATCATCTGATATTTTCTGCTTTTATGCTGACACCTCTT 1073  
 Db 901 CTGCTTTTCTTCTTATGAGTATCATCTGATATTTTCTGCTTTTATGCTGACACCTCTT 960

QY 1074 TTTAAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTGTACTGTGGCTTTTGGATTT 1133  
 Db 961 TTTAAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTGTACTGTGGCTTTTGGATTT 1020

QY 1134 ATTGCCCTTATGATAATCCTCATAGAAAGTTTCCCAAAATCGTTAGTGTGGCTTTTCAGT 1193  
DB 1021 ATTGCCCTTATGATAATCCTCATAGAAAGTTTCCCAAAATCGTTAGTGTGGCTTTTCAGT 1080  
QY 1194 COTTTCTGTCTACTGTACTTGTGATTGCGTATTCGACAGGTCATGCACTTTAGAGATTTT 1253  
DB 1081 COTTTCTGTCTACTGTACTTGTGATTGCGTATTCGACAGGTCATGCACTTTAGAGATTTT 1140  
QY 1254 AATGAAGTGTCTTCAATTTCAAAATTTGACTCGAGGCCCATATCTCTTAATTTATTTACAAAT 1313  
DB 1141 AATGAAGTGTCTTCAATTTCAAAATTTGACTCGAGGCCCATATCTCTTAATTTATTTACAAAT 1200  
QY 1314 ATCATGCTCAGACCTTAATAGTATATCTATGTCCTCTTGCGCTGTCTATCTGTATCAAGTC 1373  
DB 1201 ATCATGCTCAGACCTTAATAGTATATCTATGTCCTCTTGCGCTGTCTATCTGTATCAAGTC 1260  
QY 1374 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTTATATTTTCTGAAGCCTTCATATTGG 1433  
DB 1261 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTTATATTTTCTGAAGCCTTCATATTGG 1320  
QY 1434 TCAAGAGCAAAAGAAATATGAGGAGTTATCAGGGCAATGTATGGAATTTATAGT 1493  
DB 1321 TCAAGAGCAAAAGAAATATGAGGAGTTATCAGGGCAATGTATGGAATTTATAGT 1380  
QY 1494 TTTAGTGAATTTATTGAGCCAGTTTCTTTCAGAAATTTGTAGAAAAGAGCCATAGAAT 1553  
DB 1381 TTTAGTGAATTTATTGAGCCAGTTTCTTTCAGAAATTTGTAGAAAAGAGCCATAGAAT 1440  
QY 1554 AGTGTATTCAGAGACATACAGAAAGAGGGTGAAATGTGGAGGCTTTGAGAAATTTG 1613  
DB 1441 AGTGTATTCAGAGACATACAGAAAGAGGGTGAAATGTGGAGGCTTTGAGAAATTTG 1500  
QY 1614 TCATTTGACATATAGCGGTCAGATTACTGCTTACTTTGCGCCACAGTGGAAACAGGAAAG 1673  
DB 1501 TCATTTGACATATAGCGGTCAGATTACTGCTTACTTTGCGCCACAGTGGAAACAGGAAAG 1560  
QY 1674 AGTACATTTGATGAAATTTCTTTGTGGACTGTGCCACACCTTCTGATGGGTTTGCATCTATA 1733  
DB 1561 AGTACATTTGATGAAATTTCTTTGTGGACTGTGCCACACCTTCTGATGGGTTTGCATCTATA 1620  
QY 1734 TATGACACAGAGTCTCAGAAATAGATCAATGTTTGAAGCAAGAAATGATTTGGCATT 1793  
DB 1621 TATGACACAGAGTCTCAGAAATAGATCAATGTTTGAAGCAAGAAATGATTTGGCATT 1680  
QY 1794 TGTCCACAGTTAGATATACATTTTGATGTTTGTGACGTAGAGAAATTTATCAATTTTG 1853  
DB 1681 TGTCCACAGTTAGATATACATTTTGATGTTTGTGACGTAGAGAAATTTATCAATTTTG 1740  
QY 1854 GCTTCAATCAAGGATACCGCCAAATATATACAGAGTGCAGAGGTTTACTA 1913  
DB 1741 GCTTCAATCAAGGATACCGCCAAATATATACAGAGTGCAGAGGTTTACTA 1800  
QY 1914 GATTTAGACATGCAGACTATCAAGATTAACCAAGCTAAAAATTAAGTGGTGGTCAAAA 1973  
DB 1801 GATTTAGACATGCAGACTATCAAGATTAACCAAGCTAAAAATTAAGTGGTGGTCAAAA 1860  
QY 1974 AGAAGCTGTCTATTAGGAATTTGCTGTTCTTGGAAACCCAAAGACTCTGCTAGATGAA 2033  
DB 1861 AGAAGCTGTCTATTAGGAATTTGCTGTTCTTGGAAACCCAAAGACTCTGCTAGATGAA 1920  
QY 2034 CCAACAGCTGGAAATGGACCCCTGTTCTGACATATTTGTAATCTTTTAAAAATACAGA 2093  
DB 1921 CCAACAGCTGGAAATGGACCCCTGTTCTGACATATTTGTAATCTTTTAAAAATACAGA 1980  
QY 2094 AAAGCCAAATCGGGTGACAGTGTCTAGTACTCATTTTCATGGATGAAGCTGACATTTTGCA 2153  
DB 1981 AAAGCCAAATCGGGTGACAGTGTCTAGTACTCATTTTCATGGATGAAGCTGACATTTTGCA 2040  
QY 2154 GATAGAAAGCTGTGATATCACAGGAATGCTGGAATGTTGTTGTTCTTCAATGTTCTC 2213  
DB 2041 GATAGAAAGCTGTGATATCACAGGAATGCTGGAATGTTGTTGTTCTTCAATGTTCTC 2100  
QY 2214 AAAAGTAAATGGGGATCGGCTACCGCCTGAGCATGTATACAGACAAATATTTGTGCCACA 2273

DB 2101 AAAAGTAAATGGGGATCGGCTACCGCCTGAGCATGTATACAGACAAATTTTGTGCCACA 2160  
QY 2274 GAATCTCTTTCTTTCACTGGTTAAACCAACATATACCTTGGAGCTACTTTATTATCAACAGAAAT 2333  
DB 2161 GAATCTCTTTCTTTCACTGGTTAAACCAACATATACCTTGGAGCTACTTTATTATCAACAGAAAT 2220  
QY 2334 GACCAACAACTGTGTATAGCTTGGCTTTCAAGGACATGGAACAAATTTTCAGGTTGTTT 2393  
DB 2221 GACCAACAACTGTGTATAGCTTGGCTTTCAAGGACATGGAACAAATTTTCAGGTTGTTT 2280  
QY 2394 TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTATGGTGTGTTTCCATGACGACT 2453  
DB 2281 TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCAATTTCTTATGGGTGTTTCCATGACGACT 2340  
QY 2454 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAAATTTGACCAAGCAGATTTATAGT 2513  
DB 2341 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAAATTTGACCAAGCAGATTTATAGT 2400  
QY 2514 GTATTTTACTCAGCAGCCACTGGAGGAAGAAATTTGGAATTTCAAAATCTTTTGTGAAATGGAA 2573  
DB 2401 GTATTTTACTCAGCAGCCACTGGAGGAAGAAATTTGGAATTTCAAAATCTTTTGTGAAATGGAA 2460  
QY 2574 CAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAGCAACCATAGAGCCTTTGG 2633  
DB 2461 CAGAGCTTACTTATTTCTGAAACCAAGGCTKCTCTAGTGAGCACCATAGAGCCTTTGG 2520  
QY 2634 AAACACAGATGTATACATAGCAAGTTTCTTCTTTTACCTTGAACGTTGAAGTAA 2693  
DB 2521 AAACACAGATGTATACATAGCAAGTTTCTTCTTTTACCTTGAACGTTGAAGTAA 2580  
QY 2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2753  
DB 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2640  
QY 2754 TTGGTTTCATCACTCTTTTAAABATGCTGTGGTTCCCATCAAACTTGTTCACAGCTTATAT 2813  
DB 2641 TTGGTTTCATCACTCTTTTAAABATGCTGTGGTTCCCATCAAACTTGTTCACAGCTTATAT 2700  
QY 2814 TTTCTAAAACTGGAGACAAACCAACATAAATCAAAACAAAGTCTGCTTCTTCAAAATCT 2873  
DB 2701 TTTCTAAAACTGGAGACAAACCAACATAAATCAAAACAAAGTCTGCTTCTTCAAAATCT 2760  
QY 2874 GCTGACTCAGATATCAGTGTATTTAGCTTTTTCACAGCCAGACATATGTTGAGC 2933  
DB 2761 GCTGACTCAGATATCAGTGTATTTAGCTTTTTCACAGCCAGACATATGTTGAGC 2820  
QY 2934 ATGATTTAAATGACAGTGTATGATCCGTGGCTCCCCATAGTGGGCTTTAAATGTTGATG 2993  
DB 2821 ATGATTTAAATGACAGTGTATGATCCGTGGCTCCCCATAGTGGGCTTTAAATGTTGATG 2880  
QY 2994 CATTCAGAAAAGGACTATGTTTTTGGAGCTGTTTTCAACAGTACTATGTTTATCTTTA 3053  
DB 2881 CATTCAGAAAAGGACTATGTTTTTGGAGCTGTTTTTCAACAGTACTATGTTTATCTTTA 2940  
QY 3054 CCTATATTAGTGAATATCATTTAGTAACTACTTCTTATCAATTTAAATGTGACTGAAACC 3113  
DB 2941 CCTATATTAGTGAATATCATTTAGTAACTACTTCTTATCAATTTAAATGTGACTGAAACC 3000  
QY 3114 ATCCAGATCTGGAGTACCCCATTTCTTCAAGAAAATTTACTGATATAGTTTTTAAATTTAG 3173  
DB 3001 ATCCAGATCTGGAGTACCCCATTTCTTCAAGAAAATTTACTGATATAGTTTTTAAATTTAG 3060  
QY 3174 CTGTATTTTCAAGCAGCTTTGCTTGGAACTTTTACTGCAATGCACTTACTTTTGGC 3233  
DB 3061 CTGTATTTTCAAGCAGCTTTGCTTGGAACTTTTACTGCAATGCACTTACTTTTGGC 3120  
QY 3234 ATGGAAAATGACAGAAATCATAGATCAAAAGCTTATATCTCAACTTAAACTTTCAGGCTTT 3293  
DB 3121 ATGGAAAATGACAGAAATCATAGATCAAAAGCTTATATCTCAACTTAAACTTTCAGGCTTT 3180  
QY 3294 TTGCCATCTGCATATTGGATTGGACAGCTGTTGTGATATCCCTTATTTTATCAAT 3353

Db 3181 TTGCCATCTGCATATATGGATGGCAAGCTGTGTGTATATATCCCTTATTTTTATCATTT 3240  
 QY 3354 CTTATTTTCATGCTAGGAGCTTATGGCATTTTCATTTATGGAATATATTTTATACGTGA 3413  
 Db 3241 CTTATTTTCATGCTAGGAGCTTATGGCATTTTCATTTATGGAATATATTTTATACGTGA 3300  
 QY 3414 AAGTTCCTTTCGTGTGGTTTTTTCCTTATATGGTTATGTTCCATCAGTTATTTCTGTCACT 3473  
 Db 3301 AAGTTCCTTTCGTGTGGTTTTTTCCTTATATGGTTATGTTCCATCAGTTATTTCTGTCACT 3360  
 QY 3474 TATATTTGCTTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3533  
 Db 3361 TATATTTGCTTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
 QY 3534 TATTTCTGGGACGCTTGGCTTGTATTTGCAATCACTGAAATTAACCTTTCTTTATGGGATAC 3593  
 Db 3421 TATTTCTGGGACGCTTGGCTTGTATTTGCAATCACTGAAATTAACCTTTCTTTATGGGATAC 3480  
 QY 3594 ACAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTCATCCATCTATCCACTTCTA 3653  
 Db 3481 ACAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTCATCCATCTATCCACTTCTA 3540  
 QY 3654 GGTTCCTGCTGATTTCTTTTCAATAAGATTTCTTGGAAAGAAATGTACGAAAGAAATGTGACACC 3713  
 Db 3541 GGTTCCTGCTGATTTCTTTTCAATAAGATTTCTTGGAAAGAAATGTACGAAAGAAATGTGACACC 3600  
 QY 3714 TATATTCCTATGGATAGGCTTTTCAGTAGCTTTATATGCTTTTACCTGAGTGTACTG 3773  
 Db 3601 TATATTCCTATGGATAGGCTTTTCAGTAGCTTTATATGCTTTTACCTGAGTGTACTG 3660  
 QY 3774 TGGATTTTCTCTTCACTATCATGAAAGAAATATGAGGAGAGATCAATTAAGAAAGAT 3833  
 Db 3661 TGGATTTTCTCTTCACTATCATGAAAGAAATATGAGGAGAGATCAATTAAGAAAGAT 3720  
 QY 3834 CCTTTTTCAGAAACCTTTTCAGAGAGTCTTAAATATAGGAAGCTTCCAGAACCCAGAC 3893  
 Db 3721 CCTTTTTCAGAAACCTTTTCAGAGAGTCTTAAATATAGGAAGCTTCCAGAACCCAGAC 3780  
 QY 3894 AATGAGGATGAAGTCAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGGT 3953  
 Db 3781 AATGAGGATGAAGTCAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGGT 3840  
 QY 3954 TGCAGTGTGTGAGGAGAAACATCCATTTATGGTCAGCAATTTGCAATAAGAAATATGAT 4013  
 Db 3841 TGCAGTGTGTGAGGAGAAACATCCATTTATGGTCAGCAATTTGCAATAAGAAATATGAT 3900  
 QY 4014 GACAAGAGAGATTTCTTTTCAAGAAAGTAAAGAGTGGCACTAAATACATCTCT 4073  
 Db 3901 GACAAGAGAGATTTCTTTTCAAGAAAGTAAAGAGTGGCACTAAATACATCTCT 3960  
 QY 4074 TTCTGTGTGAAAGAGAGATCTTTAGGACTATTTGGGTCCAAATGGTGTGGCAAAAGC 4133  
 Db 3961 TTCTGTGTGAAAGAGAGATCTTTAGGACTATTTGGGTCCAAATGGTGTGGCAAAAGC 4020  
 QY 4134 ACAATTTATTAATTTCTGTTGGTGATATGACCACTTCAGGCCAGGTATTTTAGGA 4193  
 Db 4021 ACAATTTATTAATTTCTGTTGGTGATATGACCACTTCAGGCCAGGTATTTTAGGA 4080  
 QY 4194 GATTTATTTTCAGAGCAAGTGAAGTATGATTTCACTGAAGTATGGGTACTGTCTCT 4253  
 Db 4081 GATTTATTTTCAGAGCAAGTGAAGTATGATTTCACTGAAGTATGGGTACTGTCTCT 4140  
 QY 4254 CAGATTAACCTTTTGGCCAGATATCAATTTGCAAGAAACATTTTGAATTTATGAGCT 4313  
 Db 4141 CAGATTAACCTTTTGGCCAGATATCAATTTGCAAGAAACATTTTGAATTTATGAGCT 4200  
 QY 4314 GTCAAGAGGATGAGTCAAGTGCATGAAAGAGTCAATTAAGTCCGAATACATCACTT 4373  
 Db 4201 GTCAAGAGGATGAGTCAAGTGCATGAAAGAGTCAATTAAGTCCGAATACATCACTT 4260  
 QY 4374 GATTTTAAAGAAACATTTTCAGAGAGCTGTAAAGAACTACCTGCAAGGATCAAAAG 4433  
 Db 4261 GATTTTAAAGAAACATTTTCAGAGAGCTGTAAAGAACTACCTGCAAGGATCAAAAG 4320

QY 4434 TTGTGTTTGTCTCTAAGTATGCTAGGAATCCTCAGATTACTTTTGTAGATGAACCATCT 4493  
 Db 4321 TTGTGTTTGTCTCTAAGTATGCTAGGAATCCTCAGATTACTTTTGTAGATGAACCATCT 4380  
 QY 4494 ACAGGTATGATCCCAAGAACCAAGCAGCATGTGGCGAGCAATTCGAATGCAATTTAA 4553  
 Db 4381 ACAGGTATGATCCCAAGAACCAAGCAGCATGTGGCGAGCAATTCGAATGCAATTTAA 4440  
 QY 4554 AACAGAAAGCGGCTGTCTTCTTGACACTCACTATATGAGGAGGAGGAGGCTGTCTGT 4613  
 Db 4441 AACAGAAAGCGGCTGTCTTCTTGACACTCACTATATGAGGAGGAGGAGGCTGTCTGT 4500  
 QY 4614 GATCGAGTAGCTATCATGTGTCTGGCAGTTAAGATGTATCGGAACAGTACAACATCTA 4673  
 Db 4501 GATCGAGTAGCTATCATGTGTCTGGCAGTTAAGATGTATCGGAACAGTACAACATCTA 4560  
 QY 4674 AAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAGACTGGATAGAAAAC 4733  
 Db 4561 AAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAGACTGGATAGAAAAC 4620  
 QY 4734 CTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTGAG 4793  
 Db 4621 CTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTGAG 4680  
 QY 4794 GAAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGTGTTCAGTCCCTTTCA 4853  
 Db 4681 GAAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGTGTTCAGTCCCTTTCA 4740  
 QY 4854 CAATCTTTTTTAAGCTGGAAGAGCT 4880  
 Db 4741 CAATCTTTTTTAAGCTGGAAGAGCT 4767

RESULT 9  
 ADAS3416  
 ID ADAS3416 standard; cdna; 3268 BP.  
 XX AC ADAS3416;  
 XX AC AC  
 XX AC AC  
 DT 20-NOV-2003 (first entry)  
 XX Human coding sequence, SEQ ID 984.  
 DE Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;  
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;  
 KW inflammatory disease; osteoporosis; neurological disease; gene; ss.  
 XX Homo sapiens.  
 OS  
 XX  
 FN EPI293569-A2.  
 XX  
 PD 19-MAR-2003.  
 XX  
 PF 21-MAR-2002; 2002BP-00006586.  
 XX  
 PR 14-SEP-2001; 2001JP-00328381.  
 PR 24-JAN-2002; 2002US-0350435P.  
 XX  
 XX (HELI-) HELIX RES INST.  
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.  
 XX  
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX WPI; 2003-395539/38.  
 DR P-PSDB; ADA55055.  
 XX  
 XX New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.  
 XX



PS	Claim 1; SEQ ID NO 984; 205pp; English.	XX
CC	The present invention relates to novel human secretory or membrane proteins (ADA54570-ADA55710) and their coding sequences (ADA52433-ADA54071). The coding sequences are useful in the gene therapy of diseases caused by abnormalities of the proteins, e.g. cancer, inflammatory diseases, osteoporosis or neurological disease.	XX
XX	Sequence 3268 BP; 1022 A; 540 C; 622 G; 1084 T; 0 U; 0 Other;	XX
XX	Query Match 58.0%; Score 3175.8; DB 7; Length 3268; Best Local Similarity 98.8%; Pred. No. 0; Matches 3223; Conservative 0; Mismatches 7; Indels 33; Gaps 1;	XX
QY	2205 ATGTTCTCTCAAAAGTAAATGGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATAT	QY
DB	6 ATGTTCTCTCAAAAGTAAATGGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATAT	DB
QY	2265 TGTGCCACAGATCTCTTTCTTCACTGGTTAAACCAACATATACCTGGAGCTACTTTATTA	QY
DB	66 TGTGCCACAGATCTCTTTCTTCACTGGTTAAACCAACATATACCTGGAGCTACTTTATTA	DB
QY	2325 CAACAGAAATGACCAACAACTCTGTATAGCTTGCCTTTCAAGGACATGGACAAATTTTCA	QY
DB	126 CAACAGAAATGACCAACAACTCTGTATAGCTTGCCTTTCAAGGACATGGACAAATTTTCA	DB
QY	2385 G-----GTTTGTGTTCTGCCCTAGACAGTCAT	QY
DB	186 GGAATGCTGTATAGACAAAGATGTTATGATGTTGTTTCTGCCCTAGACAGTCAT	DB
QY	2412 TCAAAATTTGGGTGTCATTTCTTATGGTGTTCCTATGAGCACTTTGGAAGACGTATTTTTA	QY
DB	246 TCAAAATTTGGGTGTCATTTCTTATGGTGTTCCTATGAGCACTTTGGAAGACGTATTTTTA	DB
QY	2472 AAGCTAGAGTTGAAGCAGAAATGACCAAGCAGATTAAGTGTATTTTACCTCAGACGCCA	QY
DB	306 AAGCTAGAGTTGAAGCAGAAATGACCAAGCAGATTAAGTGTATTTTACCTCAGACGCCA	DB
QY	2532 CTGGAGGAAGAAATGAAATCAAAATCTTTTGTATGAAATGGAACAGAGCTTACTTATTCCT	QY
DB	366 CTGGAGGAAGAAATGAAATCAAAATCTTTTGTATGAAATGGAACAGAGCTTACTTATTCCT	DB
QY	2592 TCTGAAACCAAGGCTCTCTAGTGAGCACCAGAGCTTTTGGAAACCAACAGATGTATACA	QY
DB	426 TCTGAAACCAAGGCTCTCTAGTGAGCACCAGAGCTTTTGGAAACCAACAGATGTATACA	DB
QY	2652 ATAGCAAGCTTCAATTTCTTACCTTGAAACGTGAACTGAATCAGTGAGATCAGTGTG	QY
DB	486 ATAGCAAGCTTCAATTTCTTACCTTGAAACGTGAACTGAATCAGTGAGATCAGTGTG	DB
QY	2712 CTTCTGCTTTTAAATTTTTCACAGTTTATGTTTTTATGTTTTTGGTTTCATCACTCTTTT	QY
DB	546 CTTCTGCTTTTAAATTTTTCACAGTTTATGTTTTTATGTTTTTGGTTTCATCACTCTTTT	DB
QY	2772 AAAAATGCTGTGGTCCCATCAAACTGTTCCAGACTTATATTTTCTTAAACCTGGAGAC	QY
DB	606 AAAAATGCTGTGGTCCCATCAAACTGTTCCAGACTTATATTTTCTTAAACCTGGAGAC	DB
QY	2832 AAACCAATATAAACAAGCTGCTGCTTCTTCAAAATTCGTGACTCAGATATCAGT	QY
DB	666 AAACCAATATAAACAAGCTGCTGCTTCTTCAAAATTCGTGACTCAGATATCAGT	DB
QY	2892 GATCTTATTTAGCTTTTTCAGACGAGAACATTAATGTTGACGNTGATTTATGACAGTGAC	QY
DB	726 GATCTTATTTAGCTTTTTCAGACGAGAACATTAATGTTGACGNTGATTTATGACAGTGAC	DB
QY	2952 TATGTATCCGTGGCTCCCATAGTCGGCTTTAAATGTATGATGATTCAGAAAAGGACTAT	QY
DB	786 TATGTATCCGTGGCTCCCATAGTCGGCTTTAAATGTATGATGATTCAGAAAAGGACTAT	DB
QY	3012 GTTTTTGACAGCTGTTTTCAACAGTACTATGGTTTATTTCTTTACCTATATTAGTGAATATC	QY
DB	846 GTTTTTGACAGCTGTTTTCAACAGTACTATGGTTTATTTCTTTACCTATATTAGTGAATATC	DB
QY	3071 GTTTTTGACAGCTGTTTTCAACAGTACTATGGTTTATTTCTTTACCTATATTAGTGAATATC	QY
DB	905 GTTTTTGACAGCTGTTTTCAACAGTACTATGGTTTATTTCTTTACCTATATTAGTGAATATC	DB
QY	3072 ATTAGTAACACTACTCTTTTATCAATTTAAATGTGACTGAAACCATCCAGATCTGGAGTACC	QY
DB	906 ATTAGTAACACTACTCTTTTATCAITTTAAATGTGACTGAAACCATCCAGATCTGGAGTACC	DB
QY	3132 CCATTTCTTTCAAGAAATTTACTGATATAGTTTTTAAATTTGAGCTGTATTTTCAAGACAGCT	QY
DB	966 CCATTTCTTTCAAGAAATTTACTGATATAGTTTTTAAATTTGAGCTGTATTTTCAAGACAGCT	DB
QY	3192 TTGCTTTGGAATCATTTGTTTACTGCAATGCCACTTACTTTTGGCATCGGAAATGACAGAGAT	QY
DB	1026 TTGCTTTGGAATCATTTGTTTACTGCAATGCCACTTACTTTTGGCATCGGAAATGACAGAGAT	DB
QY	3252 CATTAAGATCAAAAGCTTATATCTCAACTTAAACTTTTCAAGCTTTTTTGGCCATCTGCATATGG	QY
DB	1086 CATTAAGATCAAAAGCTTATATCTCAACTTAAACTTTTCAAGCTTTTTTGGCCATCTGCATATGG	DB
QY	3312 ATTGGACAAAGCTGTTTGTGATATCCCTTATTTTTTATCATCTTATTTTGTATGCTCTAGGA	QY
DB	1146 ATTGGACAAAGCTGTTTGTGATATCCCTTATTTTTTATCATCTTATTTTGTATGCTCTAGGA	DB
QY	3372 AGCTTATTTGGCAATTTCAATTTAGGATATATTTTTATATCTGTAAGATTTCTTGTCTGGTT	QY
DB	1206 AGCTTATTTGGCAATTTCAATTTAGGATATATTTTTATATCTGTAAGATTTCTTGTCTGGTT	DB
QY	3432 TTTTGGCTTATTTGGTTATGTTTCCATCAGTTATTTCTGTTTCACTTATATGCTTCTTCCACC	QY
DB	1266 TTTTGGCTTATTTGGTTATGTTTCCATCAGTTATTTCTGTTTCACTTATATGCTTCTTCCACC	DB
QY	3492 TTTTAAAGAAATTTTAAATPACCAAGAAATTTTGGTCAATTTATCTATCTGTTGGCAGCGTTG	QY
DB	1326 TTTTAAAGAAATTTTAAATPACCAAGAAATTTTGGTCAATTTATCTATCTGTTGGCAGCGTTG	DB
QY	3552 GCTTGTATGCAATGCACTGAAATTAACCTTTCTTTATGGGATACACAATGCAATATTTCTT	QY
DB	1386 GCTTGTATGCAATGCACTGAAATTAACCTTTCTTTATGGGATACACAATGCAATATTTCTT	DB
QY	3612 CAATATGCTTTTGTATCATCATTTCCAATCTATCCACTTCTAGGTTGCCGTGATTTCTTTC	QY
DB	1446 CAATATGCTTTTGTATCATCATTTCCAATCTATCCACTTCTAGGTTGCCGTGATTTCTTTC	DB
QY	3672 ATAAAGATTTCTTGAAGAATGTACGAAATAATTTGACACCTTATATCCATGCGGATAGG	QY
DB	1506 ATAAAGATTTCTTGAAGAATGTACGAAATAATTTGACACCTTATATCCATGCGGATAGG	DB
QY	3732 CTTTTCAGTAGCTGTTTATATCGCTTACCTGCAGTGTCTGCTGCGATTTTCTCTTCAAA	QY
DB	1566 CTTTTCAGTAGCTGTTTATATCGCTTACCTGCAGTGTCTGCTGCGATTTTCTCTTCAAA	DB
QY	3792 TACTATGAGAAAAAATATGAGGAGCAGATCAATAAGAAAAAGATCCCTTTTTCAGAAAACCTT	QY
DB	1626 TACTATGAGAAAAAATATGAGGAGCAGATCAATAAGAAAAAGATCCCTTTTTCAGAAAACCTT	DB
QY	3852 TCAACGGAAGTCTAAAAATAGGAAGCTTCCAGAACCCAGACCAATGAGGATGAAGATGAA	QY
DB	1686 TCAACGGAAGTCTAAAAATAGGAAGCTTCCAGAACCCAGACCAATGAGGATGAAGATGAA	DB
QY	3912 GATGTCAAAGCTGAAAAGCTAAAAGCTCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAG	QY
DB	1746 GATGTCAAAGCTGAAAAGCTAAAAGCTCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAG	DB
QY	3972 AAACCATCCATATGCTCAGCAATTTGCATTAAGAAATATGATGACAGAAAGATTTTCTT	QY
DB	1806 AAACCATCCATAT	



Db 61 GCACCATGAGCCCTTTGGAAACCAACAGATGTATACAAATGCAAGATTTTCATTTCTTTACCT 120  
Qy 2677 TGAACGCGTGAAGTAAATCAGTGAGATCAGTGTCTGCTCTGCTTTTAAATTTTTTTTACAG 2736  
Db 121 TGAACGCGTGAAGTAAATCAGTGAGATCAGTGTCTGCTTTTAAATTTTTTTTACAG 180  
Qy 2737 TTCAGATTTTATGTTTTGTTGTTTCATCAGTCTTTTAAATAATGCTGTGTTCCCATCAAC 2796  
Db 181 TTCAGATTTTATGTTTTGTTGTTTCATCAGTCTTTTAAATAATGCTGTGTTCCCATCAAC 240  
Qy 2797 TTGTTCCAGACTTATATTTTCTAAACCTTGGAGACAAACCAATATAATACAAACCAAGTC 2856  
Db 241 TTGTTCCAGACTTATATTTTCTAAACCTTGGAGACAAACCAATATAATACAAACCAAGTC 300  
Qy 2857 TGCCTTTTCAAAATCTGCTGATCAGATATCAGTGATCTTTATAGCTTTTTCACAGCC 2916  
Db 301 TGCCTTTTCAAAATCTGCTGATCAGATATCAGTGATCTTTATAGCTTTTTCACAGCC 360  
Qy 2917 AGAACATAATGCTGACGATGATTAATGACAGTGACTATGATCCGTGGCTCCCATAGTG 2976  
Db 361 AGAACATAATGCTGACGATGATTAATGACAGTGACTATGATCCGTGGCTCCCATAGTG 420  
Qy 2977 CGGCTTTAAATGATGATGATTTAGAAAAGGACTATGTTTTTTCAGACTGTTTTCAACAGTA 3036  
Db 421 CGGCTTTAAATGATGATGATTTAGAAAAGGACTATGTTTTTTCAGACTGTTTTCAACAGTA 480  
Qy 3037 CTATGTTTTATTTCTTTTACCTATATTTAGTGAATATCATTTAGTAATCTATCTTTATCATTT 3096  
Db 481 CTATGTTTTATTTCTTTTACCTATATTTAGTGAATATCATTTAGTAATCTATCTTTATCATTT 540  
Qy 3097 TAAATGTGACTGAACCAACCCAGATCTGGAGTACCCCAATCTTTTCAAGAAATTAAGTATA 3156  
Db 541 TAAATGTGACTGAACCAACCCAGATCTGGAGTACCCCAATCTTTTCAAGAAATTAAGTATA 600  
Qy 3157 TAGTTTTTAAATTCAGCTGATTTTCAAGAGCTTTCTGCGATCATTTGTTACTGCA 3216  
Db 601 TAGTTTTTAAATTCAGCTGATTTTCAAGAGCTTTCTGCGATCATTTGTTACTGCA 660  
Qy 3217 TCCACCTTTACTTTGCCATGGAATGAGAGATCATTAAGATCAAAAGCTTTACTCAAC 3276  
Db 661 TCCACCTTTACTTTGCCATGGAATGAGAGATCATTAAGATCAAAAGCTTTACTCAAC 720  
Qy 3277 TTAATACTTCAGCTTTTGGCATCTGCATATGGAATGGAAGCTGTTGTTGATATCC 3336  
Db 721 TTAATACTTCAGCTTTTGGCATCTGCATATGGAATGGAAGCTGTTGTTGATATCC 780  
Qy 3337 CCTTATTTTTTATCATTTCTTATTTTGTATGCTAGGAAGCTTTATGCAATTTTCAATATGAT 3396  
Db 781 CCTTATTTTTTATCATTTCTTATTTTGTATGCTAGGAAGCTTTATGCAATTTTCAATATGAT 840  
Qy 3397 TATATTTTTTATCTGTAAAGTTCTTGTGTTGTTTGTGCTTTTGGCTTTATTTGGTTATGTTCCAT 3456  
Db 841 TATATTTTTTATCTGTAAAGTTCTTGTGTTGTTTGTGCTTTTGGCTTTATTTGGTTATGTTCCAT 900  
Qy 3457 CAGTTTATTTCTGTTCACTTATATGCTTTTCACTTTTAAAGAAATTTTAAATACCAAG 3516  
Db 901 CAGTTTATTTCTGTTCACTTATATGCTTTTCACTTTTAAAGAAATTTTAAATACCAAG 960  
Qy 3517 AATTTTGTCTATTTATCTATTTCTGTCGAGCGTTGGCTTTGTTATTTGCAATCACTGAAATAA 3576  
Db 961 AATTTTGTCTATTTATCTATTTCTGTCGAGCGTTGGCTTTGTTATTTGCAATCACTGAAATAA 1020  
Qy 3577 CTTTCTTTTATGGATACCAATTCGAATTTCTTCAATTTGATGCTTTTGTATCATCTTC 3636  
Db 1021 CTTTCTTTTATGGATACCAATTCGAATTTCTTCAATTTGATGCTTTTGTATCATCTTC 1080  
Qy 3637 CAATCTATCCACTTCTAGGTTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGATC 3696  
Db 1081 CAATCTATCCACTTCTAGGTTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGATC 1140  
Qy 3697 GAAATAATGTGACACCTTAAATCCATGGGATAGCTTTTCTAGTACTGTTTATATCGCCTT 3756  
Db 1141 GAAATAATGTGACACCTTAAATCCATGGGATAGCTTTTCTAGTACTGTTTATATATCGCCTT 1200

Qy 3757 ACCTGAGTGTGACTGTGATTTTCTCTTACAATACTATGAGAAAAATATGGAGCA 3816  
Db 1201 ACCTGAGTGTGACTGTGATTTTCTCTTACAATACTATGAGAAAAATATGGAGCA 1260  
Qy 3817 GATCAATAAGAAAAAGATCCCTTTTCAGAAACCTTTCAAGAAAGTCTTAAATAATGGAAGC 3876  
Db 1261 GATCAATAAGAAAAAGATCCCTTTTCAGAAACCTTTCAAGAAAGTCTTAAATAATGGAAGC 1320  
Qy 3877 TTCAGAACCCACAGACAATGAGGATGAAGATGCAAGATGCAAGCTGAAAGACTAAAGG 3936  
Db 1321 TTCAGAACCCACAGACAATGAGGATGAAGATGCAAGATGCAAGCTGAAAGACTAAAGG 1380  
Qy 3937 TCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAAACCATCTATATGGTCAGCAAT 3996  
Db 1381 TCAAGAGCTGATGGGTTGCCAGTGTGTGAGGAGAAAACCATCTATATGGTCAGCAAT 1440  
Qy 3997 TGCATAAAGAAATATGATGACAAGAAAGATTTCTTCTTCAAGAAAGTAAAGAAAGTGG 4056  
Db 1441 TGCATAAAGAAATATGATGACAAGAAAGATTTCTTCTTCAAGAAAGTAAAGAAAGTGG 1500  
Qy 4057 CAACATAATACATCTCTTTCTGTGTGAAAAGAGAGAGATCTTTAGCACTATTTGGGTCAA 4116  
Db 1501 CAACATAATACATCTCTTTCTGTGTGAAAAGAGAGAGATCTTTAGCACTATTTGGGTCAA 1560  
Qy 4117 ATGGTCTGGCAAGACAAATTTATTAATATTTCTGTTGGTGATATTTGAACCAACTCAG 4176  
Db 1561 ATGGTCTGGCAAGACAAATTTATTAATTTCTGTTGGTGATATTTGAACCAACTCAG 1620  
Qy 4177 GCCAGGTATTTTTAGGAGATTTATTTCTCAGAGACAAGTGAAGATGATGATTTCACTGAAGT 4236  
Db 1621 GCCAGGTATTTTTAGGAGATTTATTTCTCAGAGACAAGTGAAGATGATGATTTCACTGAAGT 1680  
Qy 4237 GTATGGTTACTGCTCTCAGATAAACCTTTGTGGCCAGATATCTACATTTGCGAGAACAT 4296  
Db 1681 GTATGGTTACTGCTCTCAGATAAACCTTTGTGGCCAGATATCTACATTTGCGAGAACAT 1740  
Qy 4297 TTAGAAATTTATGGAGCTGTCAAAGGAATGAGTGCAGTGCATGAAAGAAAGTCAATAGTC 4356  
Db 1741 TTAGAAATTTATGGAGCTGTCAAAGGAATGAGTGCAGTGCATGAAAGAAAGTCAATAGTC 1800  
Qy 4357 GAATAACATGACATTTGATTTTAAAGAACATCTTTCAGAAAGCTGTAAAGAAACTACCTG 4416  
Db 1801 GAATAACATGACATTTGATTTTAAAGAACATCTTTCAGAAAGCTGTAAAGAAACTACCTG 1860  
Qy 4417 CAGGAATCAAAAGAAAGTTGTTTTGCTCTAAGTATGCTAGGAACTCTCAGATTTACTT 4476  
Db 1861 CAGGAATCAAAAGAAAGTTGTTTTGCTCTAAGTATGCTAGGAACTCTCAGATTTACTT 1920  
Qy 4477 TGTAGATGAACCATCTACAGGTATGGATCCCAAAGCCAAAACAGCACATGTGGCGAGCAA 4536  
Db 1921 TGTAGATGAACCATCTACAGGTATGGATCCCAAAGCCAAAACAGCACATGTGGCGAGCAA 1980  
Qy 4537 TTCGAACCTCATTTAAAAACAGAAACGGGCTGCTATTTCTGACCACTCACTATATGGAGG 4596  
Db 1981 TTCGAACCTCATTTAAAAACAGAAAGCGGCTGCTATTTCTGACCACTCACTATATGGAGG 2040  
Qy 4597 AGCGAGAGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGCGAGTTTAAAGATGTATCG 4656  
Db 2041 AGCGAGAGCTGTCTGTGATCGAGTAGCTATCATGTGTCTGGCGAGTTTAAAGATGTATCG 2100  
Qy 4657 GAACAGTACAACATCTTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTAATATGA 4716  
Db 2101 GAACAGTACAACATCTTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTAATATGA 2160  
Qy 4717 AGGACTGGATGAAAACTTAGAAGTAGACCGCTTTCAAGAGAAATTCAGTATATTTTCC 4776  
Db 2161 AGGACTGGATGAAAACTTAGAAGTAGACCGCTTTCAAGAGAAATTCAGTATATTTTCC 2220  
Qy 4777 CAAATGCAAGCGCTCAGGAAAGTTTTTCTTCTATTTTGGCTTATAAATTTCCCTAGGAG 4836  
Db 2221 CAAATGCAAGCGCTCAGGAAAGTTTTTCTTCTATTTTGGCTTATAAATTTCCCTAGGAG 2280

Qy	4837	ATGTTGAGTC	CCCTTTCTCA	CAATCTTTTTT	TAAAGCTGG	AAGCTAA	ACATGCTTT	TGGCA	4896
Db	2281	ATGTTGAGTC	CCCTTTCTCA	CAATCTTTTTT	TAAAGCTGG	AAGCTAA	ACATGCTTT	TGGCA	2340
Qy	4897	TTGAAGAAT	ATAGCTTTT	CTCAAGCA	CACATTG	GAACAGG	TTTTTGT	AGAACTC	4956
Db	2341	TTGAAGAAT	ATAGCTTTT	CTCAAGCA	CACATTG	GAACAGG	TTTTTGT	AGAACTC	2400
Qy	4957	AACAAGAGG	AGGAATAA	TAGTTGT	CGAACCTT	TAAACAG	CACACTTT	TGCTGG	5016
Db	2401	AACAAGAGG	AGGAATAA	TAGTTGT	CGAACCTT	TAAACAG	CACACTTT	TGCTGG	2460
Qy	5017	CACAAGAAG	ATAGAGTAG	TATTTG	GAATTTG	TATGTT	CGCTCTG	CTTATCTGG	5076
Db	2461	CACAAGAAG	ATAGAGTAG	TATTTG	GAATTTG	TATGTT	CGCTCTG	CTTATCTGG	2520
Qy	5077	TTCTTTTTT	CACTTAATTT	TAACTTTG	CTTTG	TAAAGGTTTTT	TATTG	GAATG	5136
Db	2521	TTCTTTTTT	CACTTAATTT	TAACTTTG	CTTTG	TAAAGGTTTTT	TATTG	GAATG	2580
Qy	5137	GAACCAAGA	ACGCACIT	TGTAATTTT	CTAAGCTC	CTTTAA	TGAAATG	CTCTG	5196
Db	2581	GAACCAAGA	ACGCACIT	TGTAATTTT	CTAAGCTC	CTTTAA	TGAAATG	CTCTG	2640
Qy	5197	TTTTTGCTT	TTCTTTAA	ATAAAA	CGTAGTAT	ATAATTA	AGTGA	AGCTGCA	5244
Db	2641	TTTTTGCTT	TTCTTTAA	ATAAAA	CGTAGTAT	ATAATTA	AGTGA	AGCTGCA	2688

RESULT 11

ABK35707	
ID	ABK35707 standard; cDNA; 3928 BP.
XX	
XX	ABK35707;
XX	
DT	08-MAY-2002 (first entry)
XX	
DE	cDNA sequence #98 encoding novel human secreted protein.
XX	
KW	Human secreted protein; hyperproliferative disorder; autoimmune disorder;
KW	immune deficiency disorder; blood disorder; inflammatory disorder;
KW	infectious disorder; allergic condition; neurodegenerative disorder;
KW	liver fibrosis; coagulation disorder; gene therapy; antimicrobial;
KW	tumour; cancer; hepatotropic; immunosuppressive; anti rheumatic; gene; ss.
XX	
CS	Homo sapiens.
XX	
PN	WC200177289-A2.
XX	
PD	18-OCT-2001.
XX	
PF	29-MAR-2001; 2001WO-US010232.
XX	
PR	05-APR-2000; 2000US-0195605P.
XX	
PA	(GEMY ) GENETICS INST INC.
PI	Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
PI	Herberg D, Treacy M, Agostino WJ, Bowman MR, Spaulding V, Wong GG;
PI	Clark HF, Fechtel K, Howes SH, Resnick RJ, Gulukota K, Graham JR;
XX	
DR	WPI; 2002-179322/23.
XX	
PT	Six hundred and twenty three polynucleotides derived from a variety of
PT	human tissue sources which encode secreted proteins, useful for treating
PT	immune deficiencies and disorders such as autoimmune disorders.
XX	
PS	Claim 1; Page 127-128; 393pp; English.
XX	
CC	The present invention relates to the isolation of novel cDNA sequences
CC	which encode human secreted proteins. The cDNA sequences have been
CC	derived from a variety of human tissues. The invention also provides a
CC	method for producing proteins from these polynucleotide sequences. The

Db 781 TTGACAGTAGAAGAAATTTATCAATTTTGGCTTCAATCAAGGGATACAGCCAAAT 840  
 QY 1884 ATAATACAGAGAGTGCAGAGGTTTTTACTAGATTTTAGCATGCGACATCATCAAGATAAC 1943  
 Db 841 ATAATACAGAGAGTGCAGAGGTTTTTACTAGATTTTAGCATGCGACATCATCAAGATAAC 900  
 QY 1944 CAAGCTAAATAATTAAGTGGTGGTCAAAAAGAAAGCTGTCTATTAGGAATTTGCTTTCTT 2003  
 Db 901 CAAGCTAAATAATTAAGTGGTGGTCAAAAAGAAAGCTGTCTATTAGGAATTTGCTTTCTT 960  
 QY 2004 GGGACCCCAAGATCTGCTGTAGATGAACCAACAGCTGGATGGAGCCCTGTCTCGA 2063  
 Db 961 GGGACCCCAAGATCTGCTGTAGATGAACCAACAGCTGGATGGAGCCCTGTCTCGA 1020  
 QY 2064 CATATTGTATGAATCTTTTAAATACAGAAAAGCAATCGGCTGACAGTGTTCAGTACT 2123  
 Db 1021 CATATTGTATGAATCTTTTAAATACAGAAAAGCAATCGGCTGACAGTGTTCAGTACT 1080  
 QY 2124 CATTTTCATGGATGAAGCTGACATCTTTCAGATAGGAAGCTGTGATATCAAGGAATG 2183  
 Db 1081 CATTTTCATGGATGAAGCTGACATCTTTCAGATAGGAAGCTGTGATATCAAGGAATG 1140  
 QY 2184 CTGAATGTGTGGTCTTCTCAATGTTCTCAAAAGTAAATGGGATCGGCTACCGCTG 2243  
 Db 1141 CTGAATGTGTGGTCTTCTCAATGTTCTCAAAAGTAAATGGGATCGGCTACCGCTG 1200  
 QY 2244 AGCATGTATAGACAAATATTTGTCACAGAAATCTTTCTTCACTGGTTTAAACAAT 2303  
 Db 1201 AGCATGTATAGACAAATATTTGTCACAGAAATCTTTCTTCACTGGTTTAAACAAT 1260  
 QY 2304 ATACTCTGAGCTACTTTTATCAACAGAAATGACCAACACTTGTATAGCTTCCCTTC 2363  
 Db 1261 ATACTCTGAGCTACTTTTATCAACAGAAATGACCAACACTTGTATAGCTTCCCTTC 1320  
 QY 2364 AAGACATGAGCAAAATTTTCAG-----GTTG 2390  
 Db 1321 AAGACATGAGCAAAATTTTCAGAAATGCTTGTATAGCAAAAGATGTTATGATGTTG 1380  
 QY 2391 TTTTCTGCTGAGCAGTCAATCAAAATTTGGTGTCTATTTCTTATGTTTTCATGACG 2450  
 Db 1381 TTTTCTGCTGAGCAGTCAATCAAAATTTGGTGTCTATTTCTTATGTTTTCATGACG 1440  
 QY 2451 ACTTTGGAAGAGCTATTTTAAAGTAGAAGTTGAAGCAGAAATTTGACCAAGCAGATTAT 2510  
 Db 1441 ACTTTGGAAGAGCTATTTTAAAGTAGAAGTTGAAGCAGAAATTTGACCAAGCAGATTAT 1500  
 QY 2511 AGTGTATTTACTCAGCAGCCACTGGAGAGAAATGGATTCAAAATCTTTTGTGAAATG 2570  
 Db 1501 AGTGTATTTACTCAGCAGCCACTGGAGAGAAATGGATTCAAAATCTTTTGTGAAATG 1560  
 QY 2571 GAACAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTTCTAGTGAGCACCATGAGCCTT 2630  
 Db 1561 GAACAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTTCTAGTGAGCACCATGAGCCTT 1620  
 QY 2631 TGGAAACACAGATGTATACATAGCAAGTTTCATTTCTTTTACCTGGAACGTGAAAGT 2690  
 Db 1621 TGGAAACACAGATGTATACATAGCAAGTTTCATTTCTTTTACCTGGAACGTGAAAGT 1680  
 QY 2691 AAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATG 2750  
 Db 1681 AAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATG 1740  
 QY 2751 TTTTGGTGTCTATCTCTTTTAAATAATCTGTGTGCTTCCATCAAACTGTGTTCCAGACTTA 2810  
 Db 1741 TTTTGGTGTCTATCTCTTTTAAATAATCTGTGTGCTTCCATCAAACTGTGTTCCAGACTTA 1800  
 QY 2811 TATTTTCTAAACCTGGAGACAAACCATATAAATACAAAACAGCTGCTTCTTCAAAAT 2870  
 Db 1801 TATTTTCTAAACCTGGAGACAAACCATATAAATACAAAACAGCTGCTTCTTCAAAAT 1860  
 QY 2871 TCTGCTG 2877  
 Db 1861 TCTGCTG 1867

RESULT 12

ABK35706  
 ID ABK35706 standard; cDNA; 1548 BP.

AC AC  
 XX ABK35706;

DT 08-MAY-2002 (first entry)

XX cDNA sequence #97 encoding novel human secreted protein.

XX Human secreted protein; hyperproliferative disorder; autoimmune disorder;  
 XX immune deficiency disorder; blood disorder; inflammatory disorder;  
 KW infectious disorder; allergic condition; neurodegenerative disorder;  
 KW liver fibrosis; coagulation disorder; gene therapy; antimicrobial;  
 KW tumour; cancer; hepatotropic; immunosuppressive; antirheumatic; gene; ss.

OS Homo sapiens.

XX WO200177289-A2.

PD 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US010232.

XX 06-APR-2000; 2000US-0195605P.

XX (GEMY ) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;  
 PI Merberg D, Treacy M, Agostino MJ, Bowman MR, Spaulding V, Wong GG;  
 PI Clark HP, Fectel K, Howes SH, Resnick RJ, Gulukota K, Graham JR;

XX WPI; 2002-179322/23.

XX Six hundred and twenty three polynucleotides derived from a variety of  
 PT human tissue sources which encode secreted proteins, useful for treating  
 PT immune deficiencies and disorders such as autoimmune disorders.

PS Claim 1; Page 126; 393pp; English.

XX The present invention relates to the isolation of novel cDNA sequences  
 CC which encode human secreted proteins. The cDNA sequences have been  
 CC derived from a variety of human tissues. The invention also provides a  
 CC method for producing proteins from these polynucleotide sequences. The  
 CC proteins are useful for identifying compounds that modulate their  
 CC activity and production. The sequences of the invention are useful for  
 CC treating diseases such as hyperproliferative disorders (e.g. cancer),  
 CC immune deficiency disorders (e.g. severe combined immunodeficiency  
 CC (SCID), autoimmune disorders (e.g. multiple sclerosis), blood disorders  
 CC (e.g. thrombocytopaenia), inflammatory disorders (e.g. arthritis),  
 CC infectious disorders (e.g. hepatitis), allergic conditions (e.g. asthma),  
 CC neurodegenerative disorders (e.g. Alzheimer's disease), liver fibrosis,  
 CC coagulation disorders (e.g. haemophilia), and tumours. The polynucleotide  
 CC sequences of the invention are also useful in gene therapy. ABK35610-  
 CC ABK36232 represent the cDNA sequences of the invention that encode for  
 CC novel human secreted proteins

XX Sequence 1548 BP; 466 A; 273 C; 268 G; 541 T; 0 U; 0 Other;

Query Match 28.2%; Score 1546.4; DB 6; Length 1548;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1547; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2489 AGAAATTGACCAAGCAGATTATAGTGTATTTACTCAGCAGCCACTGGAGAGAAATGGA 2548

Db 1 AGAATTGACCAAGCAGATTATAGTGTATTTACTCAGCAGCCACTGGAGAGAAATGGA 60

QY 2549 TTCAAAATCTTTGTATGAATGGAACAGAGCTTACTTTCTTTGAAACCAAGGCTTC 2608

Db 61 TTCAAAATCTTTGTATGAATGGAACAGAGCTTACTTTCTTTGAAACCAAGGCTTC 120



PT neurological disorders (e.g. Alzheimer's), or cell proliferations or  
 PT cancers.  
 XX  
 PS Claim 1; SEQ ID NO 102; 458bp + Sequence Listing; English.  
 XX  
 CC The invention relates to a secretory polynucleotide (designated sptm)  
 CC comprising any of 567 polynucleotide sequences (AB235837-AB236403), a  
 CC naturally occurring polynucleotide sequence at least 90 % identical to  
 CC the polynucleotide sequence, a polynucleotide complementary to them or an  
 CC RNA equivalent of them. The polypeptide or polynucleotide are useful for  
 CC treating, preventing or diagnosing a disease or condition associated with  
 CC the expression of functional SPTM. These are particularly useful for  
 CC diagnosing, treating or preventing autoimmune/inflammatory disorders  
 CC (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's  
 CC disease), neurological disorders (e.g. epilepsy, Huntington's disease,  
 CC dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,  
 CC multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,  
 CC schizophrenia or amnesia), or cell proliferative disorders (e.g.  
 CC psoriasis, polycythemia vera, or cancers including adenocarcinoma,  
 CC leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,  
 CC breast, cervix or prostate). Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences)  
 XX  
 SQ Sequence 2481 BP; 662 A; 452 C; 477 G; 889 T; 0 U; 1 Other;  
 Query Match 25.6%; Score 1400; DB 7; Length 2481;  
 Best Local Similarity 99.9%; Pred. No. 2.7e-295;  
 Matches 1400; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 97 AGGTTTATTCAGAAACATGTCACCTGCAATTAGGAGGTAGGATTTGGAGACGACCA 156  
 DB 1081 AGGTTTATTCAGAAACATGTCACCTGCAATTAGGAGGTAGGATTTGGAGACGACCA 1140  
 QY 157 GAACACTTCTACTGAAGATTACTTAATTAATGAGAACCAAAAGAGTAGTGTTCAGG 216  
 DB 1141 GAACACTTCTACTGAAGATTACTTAATTAATGAGAACCAAAAGAGTAGTGTTCAGG 1200  
 QY 217 AAATCTCTTTTCCACTATTTTATTTTGGTTAATTAATTAAGCAATGATGATCAAA 276  
 DB 1201 AAATCTCTTTTCCACTATTTTATTTTGGTTAATTAATTAAGCAATGATGATCAAA 1260  
 QY 277 ATAAGAAATATGAGAGGTCCTAATATAGACTCAATCCCTATGACAGAGTTTACTCTTT 336  
 DB 1261 ATAAGAAATATGAGAGGTCCTAATATAGAACTCAATCCCTATGACAGAGTTTACTCTTT 1320  
 QY 337 CTAATCTAATCTTGGATATCTCCAGTGACTAATATTAACAGCAGCATCATGCGAAG 396  
 DB 1321 CTAATCTAATCTTGGATATCTCCAGTGACTAATATTAACAGCAGCATCATGCGAAG 1380  
 QY 397 TGTCTACTGATCATCTACCTGATGTCATTAATTAATGAGAAATATACAAATGAAAGAA 456  
 DB 1381 TGTCTACTGATCATCTACCTGATGTCATTAATTAATGAGAAATATACAAATGAAAGAA 1440  
 QY 457 TGTAAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAGACTCCA 516  
 DB 1441 TGTAAACATCCAGTCTCTCTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAGACTCCA 1500  
 QY 517 TGTCTCTATGAATCTGTTTTTCTGATATGATTAATTAATTAATTAATTAATTAATTA 576  
 DB 1501 TGTCTCTATGAATCTGTTTTTCTGATATGATTAATTAATTAATTAATTAATTAATTA 1560  
 QY 577 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTCCTCAGGTTTCACAG 636  
 DB 1561 CAAGAGCTGGCTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTCCTCAGGTTTCACAG 1620  
 QY 637 TTTTACAAGCATCCATAGATGTCGCAATTAATACAGTTGAAGCAAAATGTTCTCTTTGGA 696  
 DB 1621 TTTTACAAGCATCCATAGATGTCGCAATTAATACAGTTGAAGCAAAATGTTCTCTTTGGA 1680  
 QY 697 AGGAGCTGGAGTCAACTAAAGCTGTTTATTTAGGGAGAACTGCTGTTGAGAAATAGATA 756  
 DB 1681 AGGAGCTGGAGTCAACTAAAGCTGTTTATTTAGGGAGAACTGCTGTTGAGAAATAGATA 1740

QY 757 CCTTCCCGGAGGAGTAATTTTAAATATATACCTAGTATATAGCAATTTTACCTTTGGATCT 816  
 DB 1741 CCTTCCCGGAGGAGTAATTTTAAATATATACCTAGTATATAGCAATTTTACCTTTGGATCT 1800  
 QY 817 TTTTGGCAATTCATATCGTAGCAGAAAGAAAAAATAAAGAAATTTTAAAGATAA 876  
 DB 1801 TTTTGGCAATTCATATCGTAGCAGAAAGAAAAAATAAAGAAATTTTAAAGATAA 1860  
 QY 877 TGGGACTTCATGATACGCTTTTGGCTTTCTCGGGTTCTTCTATATACAAGTTTAAATTT 936  
 DB 1861 TGGGACTTCATGATACGCTTTTGGCTTTCTCGGGTTCTTCTATATACAAGTTTAAATTT 1920  
 QY 937 TTCTTATGTCCTTCTTATGCGCAGTTCATGCGCAGCTTCTTGTGTTATTTCCCTCAAAGTA 996  
 DB 1921 TTCTTATGTCCTTCTTATGCGCAGTTCATGCGCAGCTTCTTGTGTTATTTCCCTCAAAGTA 1980  
 QY 997 GCAGCATTTGATATTTCTGCTTTTCTTTTATGAGTATCATCTGATTTTGTGCTT 1056  
 DB 1981 GCAGCATTTGATATTTCTGCTTTTCTTTTATGAGTATCATCTGATTTTGTGCTT 2040  
 QY 1057 TAATGCTGACACCTCTTTTAAATAATCAAAACATGTGGGAATGTTGAATTTTGTGTTA 1116  
 DB 2041 TAATGCTGACACCTCTTTTAAATAATCAAAACATGTGGGAATGTTGAATTTTGTGTTA 2100  
 QY 1117 CTGTGGCTTTTGGATTTATTTGGCTTTATGATAATCTCTATAGAAAGTTTCCCAAATCGT 1176  
 DB 2101 CTGTGGCTTTTGGATTTATTTGGCTTTATGATAATCTCTATAGAAAGTTTCCCAAATCGT 2160  
 QY 1177 TAGTGTGGCTTTTCACTCTTTCTGTCACGTACTTTTGTGATTTGATTTGCACAGTCA 1236  
 DB 2161 TAGTGTGGCTTTTCACTCTTTCTGTCACGTACTTTTGTGATTTGATTTGCACAGTCA 2220  
 QY 1237 TGCAATTTAGAAAGTTTAAATGAAGTGTCTTCAATTTTCAAAATTTGACTGAGGCCCATATC 1296  
 DB 2221 TGCAATTTAGAAAGTTTAAATGAAGTGTCTTCAATTTTCAAAATTTGACTGAGGCCCATATC 2280  
 QY 1297 CTCTAATTTAGAAATTCATCTCCTCACACTTAATAGTATTTCTATGCTCTTGGCTG 1356  
 DB 2281 CTCTAATTTAGAAATTCATCTCCTCACACTTAATAGTATTTCTATGCTCTTGGCTG 2340  
 QY 1357 TCTATCTTGATCAAGTCATTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATTTTC 1416  
 DB 2341 TCTATCTTGATCAAGTCATTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATTTTC 2400  
 QY 1417 TGAAGCTTTCATATTTGGTCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCAATG 1476  
 DB 2401 TGAAGCTTTCATATTTGGTCAAGAGCAAAAGAAATTTATGAGGATTTATCAGAGGCAATG 2460  
 QY 1477 TTAATGGAATATTTAGTTTGA 1497  
 DB 2461 TTAATGGAATATTTAGTTTGA 2481  
 RESULT 14  
 ABL64459  
 ID ABL64459 standard; DNA; 1346 BP.  
 XX  
 AC ABL64459;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Stomach cancer related gene sequence SEQ ID NO:2796.  
 XX  
 KW Human; cancer; colon; breast; ovary; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumor; cancerous;  
 KW cytosolic; gene therapy; antineoplastic; Wilms' tumor; adenocarcinoma;  
 XX gene; ds.  
 OS Homo sapiens.  
 XX  
 PN W0200194629-A2.  
 XX



PD	13-DEC-2001.	CC	prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell
XX		CC	cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous
PF	30-MAY-2001; 2001WO-US010838.	CC	cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilms
XX		CC	tumour
PR	05-JUN-2000; 2000US-0209473P.	QQ	Sequence 1346 BP; 451 A; 206 C; 290 G; 395 T; 0 U; 4 Other;
PR	05-JUN-2000; 2000US-0209531P.	Query Match	21.6%; Score 1182; DB 6; Length 1346;
PR	18-SEP-2000; 2000US-0233133P.	Best Local Similarity	96.5%; Pred. No. 8.3e-248;
PR	18-SEP-2000; 2000US-0233617P.	Matches 1306; Conservative	0; Mismatches 29; Indels 19; Gaps 9;
PR	20-SEP-2000; 2000US-0234009P.		
PR	20-SEP-2000; 2000US-0234034P.		
PR	20-SEP-2000; 2000US-0234052P.		
PR	22-SEP-2000; 2000US-0234509P.		
PR	22-SEP-2000; 2000US-0234567P.		
PR	22-SEP-2000; 2000US-0234923P.		
PR	25-SEP-2000; 2000US-0234924P.		
PR	25-SEP-2000; 2000US-0235077P.		
PR	25-SEP-2000; 2000US-0235082P.		
PR	25-SEP-2000; 2000US-0235134P.		
PR	25-SEP-2000; 2000US-0235280P.		
PR	26-SEP-2000; 2000US-0235637P.		
PR	26-SEP-2000; 2000US-0235638P.		
PR	27-SEP-2000; 2000US-0235711P.		
PR	27-SEP-2000; 2000US-0235720P.		
PR	27-SEP-2000; 2000US-0235840P.		
PR	27-SEP-2000; 2000US-0235863P.		
PR	28-SEP-2000; 2000US-0236028P.		
PR	28-SEP-2000; 2000US-0236032P.		
PR	28-SEP-2000; 2000US-0236033P.		
PR	28-SEP-2000; 2000US-0236034P.		
PR	28-SEP-2000; 2000US-0236109P.		
PR	28-SEP-2000; 2000US-0236111P.		
PR	29-SEP-2000; 2000US-0236842P.		
PR	29-SEP-2000; 2000US-0236891P.		
PR	02-OCT-2000; 2000US-0237172P.		
PR	02-OCT-2000; 2000US-0237173P.		
PR	02-OCT-2000; 2000US-0237278P.		
PR	02-OCT-2000; 2000US-0237294P.		
PR	02-OCT-2000; 2000US-0237295P.		
PR	02-OCT-2000; 2000US-0237316P.		
PR	03-OCT-2000; 2000US-0237425P.		
PR	03-OCT-2000; 2000US-0237598P.		
PR	03-OCT-2000; 2000US-0237604P.		
PR	03-OCT-2000; 2000US-0237606P.		
PR	03-OCT-2000; 2000US-0237608P.		
PR	01-NOV-2000; 2000US-024867P.		
PR	01-NOV-2000; 2000US-0245084P.		
XX			
PA	(AVAL-) AVALON PHARM.		
XX			
PI	Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;		
PI	Soppet DR, Weaver Z;		
XX			
XX	WPI; 2002-188264/24.		
PT	Screening for anti-neoplastic agent involves exposing cells to a chemical		
PT	agent to be tested for anti-neoplastic activity, and determining a change		
PT	in expression of a gene of a signature gene set.		
XX			
PS	Claim 1; SEQ ID NO 2796; 44pp; English.		
CC	The present invention describes a method (M1) for screening for an anti-		
CC	neoplastic agent. The method involves exposing cells to a chemical agent		
CC	to be tested for anti-neoplastic activity, determining a change in		
CC	expression of at least one gene (I) of a signature gene set, where (I)		
CC	comprises a sequence (S) selected from 8447 sequences (given in ABL61664		
CC	to ABL70110), or is at least 95% identical to (S), where a change in		
CC	expression is indicative of anti-neoplastic activity. (I) has cytostatic		
CC	activity and can be used in gene therapy. M1 can be used for screening an		
CC	anti-neoplastic agent, and can be used for producing a product which is		
CC	the data collected with respect to the anti-neoplastic agent as a result		
CC	of M1, and the data is sufficient to convey the chemical structure and/or		
CC	properties of the agent. M1 can be used in the treatment of cancer such		
CC	as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,		

Db 890 AGCGTCA-GAAAGTTTTCTTCTATTGTTGGCTTCTAAATTAATAAGGAAGATGTTGAG 948  
Qy TCCCTTTTCAAACTCTTTTAAAGTCTGGAAGAGCTAAACATGCTTTTGGCAATTGAAGAA 4904  
Db 949 TCCCTTTCCCAATCTTTTAAAGTCTGGAAGAGCTAAACATGCTTTTGGCAATTGAAGAA 1007  
Qy 4905 TATAGCTTTTCTCAAGCAATGGAACAGGTTTTGTAGAACTCACTAAAGAACAGAG 4964  
Db 1008 TATAGC-TTTTCTCAAGCAATGGAACAGGTTTTGTAGAACTCACTAAAGAACAGAG 1066  
Qy 4965 GAGGAAGATATAGTCTGGAACCTTTAAACAGCACACTTTGGTGGGAACCAACACAGAA 5024  
Db 1067 GAGGAAGATATAGTCTGGAACCTTTAAACAGCACACTTTGGTGGGAACCAACAGAA 1125  
Qy 5025 GATAGAGTAGTATTTGAAATTTGATTTGTTGGTCTGCTTACTGGGACTTCTTTCTTTT 5084  
Db 1126 GATAGAGTAGTATTTGAAATTTGATTTGTTGGTCTGCTTACTGGGACTTCTTTCTTTT 1185  
Qy 5085 CACTTAATTTTAACTTTGGTTTAAAGAGTTTTTATTGGAATGTAACCTGGAACCAAG 5144  
Db 1186 CACTTAATTTTAACTTTGGTTTAAAGAGTTTTTATTGGAATGTAACCTGGAACCAAG 1245  
Qy 5145 AACGACTTGAAATTTTCTAAGCTCCTTAATTAATGAAATGCTGGTGGTGGTGGTGGT 5204  
Db 1246 AACGACTTGAAATTTTCTAAGCTCCTTAATTAATGAAATGCTGGTGGTGGTGGTGGT 1305  
Qy 5205 TTCTTTAAATAAAGCTATGATATTAATTAAGTGA 5238  
Db 1306 TTCTTTAAATAAAGCTATGATATTAATTAAGTGA 1339

RESULT 15  
ABL63763  
ID ABL63763 standard; DNA; 1346 BP.  
XX AC ABL63763;  
XX DT 15-MAY-2002 (first entry)  
XX DE Breast cancer related gene sequence SEQ ID NO:2100.  
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytoskeletal; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
XX gene; ds.  
XX OS Homo sapiens.  
XX PN WO200194629-A2.  
XX PD 13-DEC-2001.  
XX PF 30-MAY-2001; 2001WO-US010838.  
XX PR 05-JUN-2000; 2000US-0209473P.  
PR 05-JUN-2000; 2000US-0209531P.  
PR 18-SEP-2000; 2000US-0223113P.  
PR 18-SEP-2000; 2000US-0223137P.  
PR 20-SEP-2000; 2000US-0234009P.  
PR 20-SEP-2000; 2000US-0234034P.  
PR 20-SEP-2000; 2000US-0234052P.  
PR 22-SEP-2000; 2000US-0234509P.  
PR 22-SEP-2000; 2000US-0234567P.  
PR 25-SEP-2000; 2000US-0234923P.  
PR 25-SEP-2000; 2000US-0234924P.  
PR 25-SEP-2000; 2000US-0235077P.  
PR 25-SEP-2000; 2000US-0235082P.  
PR 25-SEP-2000; 2000US-0235134P.  
PR 25-SEP-2000; 2000US-0235280P.  
PR 26-SEP-2000; 2000US-0235637P.  
PR 26-SEP-2000; 2000US-0235638P.  
PR 27-SEP-2000; 2000US-0235711P.  
PR 27-SEP-2000; 2000US-0235720P.

PR 27-SEP-2000; 2000US-0235840P.  
PR 27-SEP-2000; 2000US-0235863P.  
PR 28-SEP-2000; 2000US-0236028P.  
PR 28-SEP-2000; 2000US-0236032P.  
PR 28-SEP-2000; 2000US-0236033P.  
PR 28-SEP-2000; 2000US-0236034P.  
PR 28-SEP-2000; 2000US-0236109P.  
PR 28-SEP-2000; 2000US-0236111P.  
PR 29-SEP-2000; 2000US-0236842P.  
PR 29-SEP-2000; 2000US-0236891P.  
PR 02-OCT-2000; 2000US-0237172P.  
PR 02-OCT-2000; 2000US-0237173P.  
PR 02-OCT-2000; 2000US-0237278P.  
PR 02-OCT-2000; 2000US-0237294P.  
PR 02-OCT-2000; 2000US-0237295P.  
PR 02-OCT-2000; 2000US-0237316P.  
PR 03-OCT-2000; 2000US-0237425P.  
PR 03-OCT-2000; 2000US-0237598P.  
PR 03-OCT-2000; 2000US-0237604P.  
PR 03-OCT-2000; 2000US-0237606P.  
PR 03-OCT-2000; 2000US-0237608P.  
PR 01-NOV-2000; 2000US-0244867P.  
PR 01-NOV-2000; 2000US-0245084P.  
PA (AVAL-) AVALON PHARM.  
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX WPI; 2002-189264/24.  
XX  
XX Screening for anti-neoplastic agent involves exposing cells to a chemical  
XX agent to be tested for anti-neoplastic activity, and determining a change  
XX in expression of a gene of a signature gene set.  
XX  
XX Claim 1; SEQ ID NO 2100; 44pp; English.  
XX  
XX The present invention describes a method (M1) for screening for an anti-  
XX neoplastic agent. The method involves exposing cells to a chemical agent  
XX to be tested for anti-neoplastic activity, determining a change in  
XX expression of at least one gene (I) of a signature gene set, where (I)  
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664  
XX to ABL70110), or is at least 95% identical to (S), where a change in  
XX expression is indicative of anti-neoplastic activity. (I) has cytostatic  
XX activity and can be used in gene therapy. M1 can be used for screening an  
XX anti-neoplastic agent, and can be used for producing a product which is  
XX the data collected with respect to the anti-neoplastic agent as a result  
XX of M1, and the data is sufficient to convey the chemical structure and/or  
XX properties of the agent. M1 can be used in the treatment of cancer such  
XX as colon, breast, stomach, lung, thyroid, oesophageal, ovarian, kidney,  
XX prostate or pancreatic cancer, adenocarcinoma, carcinoma, clear cell  
XX cancer, infiltrating ductal cancer, infiltrating lobular cancer, squamous  
XX cell carcinoma, neuroendocrine carcinoma, papillary carcinoma and Wilm's  
XX tumour  
XX  
XX Sequence 1346 BP; 451 A; 206 C; 290 G; 395 T; 0 U; 4 Other;  
SQ  
Query Match 21.6%; Score 1182; DB 6; Length 1346;  
Best Local Similarity 96.5%; Pred. No. 8.3e-248;  
Matches 1306; Conservative 0; Mismatches 29; Indels 19; Gaps 9;  
Qy 3889 CAGACAATGAGGATGAAGATGAAGATGTCAGAGCTGAAGAGCTAAAGGTCAAGAGCTCA 3948  
Db 1 CAGACAATGAGGATGAAGATGAAGATGTCAGAGCTGAAGAGCTAAAGGTCAAGAGCTGA 60  
Qy 3949 TGGGTTGCCAGTGTGTGAGGAGAAACCATTCATTATGTCAGCAATTTCGATAAGAT 4008  
Db 61 TGGGTTGCCAGTGTGTGAGGAGAAACCATTCATTATGTCAGCAATTTCGATAAGAT 120  
Qy 4009 ATGATGACAAAGAAAGATTTTCTTCTTCAAGAAAAGTAAAGAAAGTGGCAACTAATACA 4068  
Db 121 ATGATGACAAAGAAAGATTTTCTTCTTCAAGAAAAGTAAAGAGAGTGGCAACTAATACA 180



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 12, 2004, 00:48:14 ; Search time 8463 Seconds

(without alignments)  
19318.851 Million cell updates/sec

Title: US-10-090-458-4

Perfect score: 5475

Sequence: 1 gcgtccggccctcgacag.....cataagcaatgtgaaagtt 5475

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estum:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vil:*
28: gb_gss1:*
29: gb_gss2:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2938.4	53.7	3950	11 AK047188	AK047188 Mus muscu
2	2811	51.3	3831	11 AK046203	AK046203 Mus muscu
3	1212.2	22.1	1529	11 AK034961	AK034961 Mus muscu
4	1182	21.6	1346	11 U66672	U66672 Homo sapien

5	837	15.3	850	14	CD250953	AGENCOURT
6	797.6	14.6	809	11	BC029426	Homo sapi
7	761.4	13.9	872	13	BQ722933	AGENCOURT
8	754.2	13.8	4620	29	AY414282	Homo sapi
9	741.4	13.5	773	13	BQ441353	AGENCOURT
10	731	13.4	747	13	BU686259	UI-CF-DU1
11	722.4	13.2	771	12	BG564439	602584385
12	719.4	13.1	793	14	CB241963	UI-CF-FNO
13	681.6	12.4	756	12	BG435856	602506942
14	676.4	12.4	697	13	BQ774317	UI-H-EZ1-
15	668	12.2	926	13	BU146839	AGENCOURT
16	661.6	12.1	728	13	BU623566	UI-H-FG1-
17	652.4	11.9	800	13	BU750944	CH3#036 H
18	641	11.7	664	14	CB044826	NISC GC06
19	637	11.6	670	12	BG570726	602591840
20	634.4	11.6	4594	29	AY414284	Mus muscu
21	633	11.6	633	14	CA773136	im60b03.Y
22	630.4	11.5	768	13	BU853701	AGENCOURT
23	628.6	11.5	724	14	CB962917	AGENCOURT
24	627	11.5	4620	29	AY414283	Pan trogl
25	625.2	11.4	694	10	BF969534	602271862
26	623	11.4	623	14	CA773849	im58a04.Y
27	621.6	11.4	3890	11	BC053340	Mus muscu
28	617	11.3	842	14	CA773183	im60a04.Y
29	612.4	11.2	614	14	CA771132	im60b03.Y
30	610.8	11.2	3357	11	BC037309	Homo sapi
31	609.6	11.1	4119	11	AK029256	Mus muscu
32	595.2	10.9	627	14	CD767029	AGENCOURT
33	589.6	10.8	898	12	BG192651	RST11768
34	586.8	10.7	961	14	CD242938	AGENCOURT
35	583.8	10.7	589	12	BG149983	im50a05
36	583.6	10.7	678	10	BE220217	im69f11.X
37	577.2	10.5	713	14	CF745072	UI-M-GV0-
38	573.6	10.5	581	14	CA941420	1r34h12.X
39	573.2	10.5	578	13	BU069128	im58a04.X
40	565.8	10.3	608	10	AW131017	xe68f06.X
41	565.2	10.3	744	9	AL700140	DRFZp686H
42	561.2	10.3	622	13	BU952395	1074a05.X
43	559.4	10.2	769	14	CD349566	UI-M-FY0-
44	559.2	10.2	3300	11	AK085555	Mus muscu
45	556.6	10.2	606	14	CD704485	EST21012

#### ALIGNMENTS

RESULT 1  
AK047188  
LOCUS AK047188 3950 bp mRNA linear HTC 20-SEP-2003  
DEFINITION Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone:B930033A02 product:ATP-BINDING CASSETTE  
PROTEIN homolog [Homo sapiens], full insert sequence.

ACCESSION AK047188

VERSION AK047188.1 GI:26338625

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus.

REFERENCE 1

Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

PUBLISHED 10349636

REFERENCE 2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

PUBLISHED 20499374

PUBLISHED 11042159

REFERENCE  
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitanai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikake, S., Inoue, K., Izawa, M., Ohara, E., Watahiki, M., Yoneda, I., Iihikawa, T., Orawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
Genome Res. 10 (11), 1757-1771 (2000)  
20530913  
11076861

REFERENCE  
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

REFERENCE  
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)  
6 (bases 1 to 3950)

REFERENCE  
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sogabe, Y., Tagami, D., Shibata, K., Shinagawa, A., Shiraki, T., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.  
Direct Submission

TITLE  
JOURNAL

Submitted (15-JUL-2001) Yoshinori Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

## COMMENT

Please visit our web site for further details.  
URL: http://genome.gsc.riken.go.jp/  
URL: http://fantom.gsc.riken.go.jp/

FEATURES  
source

Location/Qualifiers  
1..3950  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="FANTOM DB:B930033A02"  
/db\_xref="MG:2412728"  
/db\_xref="taxon:10090"  
/clone="B930033A02"  
/tissue\_type="cerebellum"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="10 days neonate"  
201..>3950  
/note="unnamed protein product; ATP-BINDING CASSETTE PROTEIN homolog [Homo sapiens] (SPTR|CAB93535, evidence: FASTQ, 89.4%ID, 76.5%length, match=3750)  
purative"  
/codon\_start=1  
/protein\_id="BAC32984.1"  
/db\_xref="GI:26338626"

## CDS

1..3950  
/translation="MATAIRDVGVWQRTLLIKYIKRCKKSKSVQELPLPLF  
WLISSMHPNRYEVSDELSPDKFSLNSVILGYTPVTNITSSIMORVCTDHLPK  
VIITEYANKEKLVASLSKSNFVGVFEDTMSYELFFPPEMIPVSSIMYNSREGCS  
KTCDAQYVSLGFTVLQASIDAAIQATNVSVMSELESTKAVIMGEAAVVIDTFPR  
MSLILVIAVATSSLPFQSSSIVIFLLIFLYLSSVFFALMTPLFKSKHVGWVFFV  
TVVGFVGLLIIVLIESFPRSLVWLPSLQCAFLGIAOVHLEDFNEGLPSLNTKEG  
PYPLITIMLADSVEYLLAVLDONVIGPBGRLRRSLYFLKPSYMSKNRYKEL  
SEGNGNTSLNIEIVPSSEFIEKAIPIISGQYKRYKTEWENALNLSFDIYEGQ  
ITALLHSGTGKSTLNLICGLCPSPDPSIYGHVSEIDEMFARKMISGQKERS  
NFDVIVENLSISLTKGIPANNIIQEVQKVLDDLMQAIKONKAKKSGGQKERS  
VGAIVGNKILLDEPTAGMDPFSRSHVNNLLKYRKANRVTVFSTHFMDEADILADR  
KAVISQMLKCGSSIFLTKSGWIGYRLSKYKIDRYCATESLSLRQHPAALLOQN  
DOQVATSLPDKMDFSGFLSALDHSNLGVSIGVSMTLEDFVLEKLEAFIDQAD  
YSVFTQQPREEEDSKSDFMQSLLILSETKASSVSMWQVQVIAKPHFUSLK  
REKSVRAVLILLIIFPAVOIEMFELHFSKNAVPIKLVPLDYFLKPGDKPHKYKTS  
LLQNSTDSNDINGLIIFPAQONIMVAMENDSPVSAAPHSALANVYREKQVPSAVF  
NSTWVCLPVMNNIISNYLYHLNVTETALOTWSTPELQETIDVFKILBYQAALLGI  
IVTAMPYFAMENAEHNIKATVQLKSLGSLPSAVGVADVDFLFFVILMLGSL  
FAHHLGIYPPAKPLAVFELIAYVFSVFTIYASFTFKILNTKEWSTFYSYVNP  
ACVAITVTFYQAVTAVFHTFCAIPIYPLGLCLISFISFKGSKNPKNENTYVNP  
DRLLVAVIMPYQCLIMIFLQHYEKIHGGRSIRKDPFFRALSKA"

Db 854 AGGGGTCATCCCTACCTACCTCGTCAATAGCCCTCTCGCCCTTCGCTACTCTCCCTGGCAAT 913  
Qy 827 TCATATCGTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCA 886  
Db 914 CCACATCGTGGCAGAAAAAGAAAGAAATTTTAAAGAAATTTTAAAGATAAATGGGACTTCA 973  
Qy 887 TGATACCTGCTTTTGGCTTCTCGGTTCTCTCTATATACAGTTTAAATTTTCTTATGTC 946  
Db 974 TGACACTGCTTTTGGCTTCTCGGTTCTCTCTATATACAGTTTAAATTTTCTTATGTC 1033  
Qy 947 CTTCTTATGGCAGTCATTCGACAGCTCTCTTGTATTTCTCTCAAGTAGCAGCATGT 1006  
Db 1034 CTTGCTTATGGCTGTCATCGCAACAGCTCTCTGTTATTTCCCTCAGAGTAGCAGCATGT 1093  
Qy 1007 GATATTTCTGCTTTTCTCTTATGATATATCATCTGATATTTTCTTAAATGCTGAC 1066  
Db 1094 GATCTTTCTACTGTTCTCTTATATGATTTGTCATCTGTTTCTGTTTCTTAAATGCTGAC 1153  
Qy 1067 ACTCTTTTAAATAATCAACATGCGGAATAGTGAATTTTGTATCTGCTGCTTT 1126  
Db 1154 GCTCTTTTAAATAATCAACATGCGGAATAGTGAATTTTGTATCTGCTGCTTT 1213  
Qy 1127 TGGATTTATGGCCTTTATGATTAATCTCTCATAGAAAGTTTCCCAATCGTTAGTGGCT 1186  
Db 1214 TGGATTTGTTGGCTGCTGATTTCTCTCATAGAAAGTTTCCCAAGTCTGCTGGTGGCT 1273  
Qy 1187 TTTTCAGTCTCTTCTGTCAGTCTACTTTTGTGATTTGTTGATTCACAGGTCTGCTTTAGA 1246  
Db 1274 CTTTCAGTCTCTTGTGTCAGTCTGCTTTCTGATTTGGGATTCACAGGTCTGCTTTAGA 1333  
Qy 1247 AGATTTTAAATGAAGTCTCTCAATTTCAAAATTTGACTGAGGCCCATATCTCTAAATAT 1306  
Db 1334 AGATTTTAAATGAAGTCTCTCAATTTCTAAATTTGACTGAGGTCTCTCTCTAAATAT 1393  
Qy 1307 TACAATTTATGCTGCACACTTAAATAGTATATTTCTATGCTCTTCTGCTGCTCTATCTTGA 1366  
Db 1394 TACTATTTATCATGCTAGCTCTTGACAGTGTGTTCTATGCTCTCTGCTGCTGTTATCTGGA 1453  
Qy 1367 TCAAGTCATTTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATATTTCTGAGCCCTTC 1426  
Db 1454 CCAAGTCATTTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAGCCATC 1513  
Qy 1427 ATATTCGTCAAGAGCAAAAGAAATTTATGAGGATTTATGAGGCAATGTTAATGGAAA 1486  
Db 1514 GTATTGCTCAAAAACAAAAGAACTATTAAGGAGCTATCGAGGGCAACATTAATGGCAA 1573  
Qy 1487 TATTAGTTTTAGTGAATTTAGCCAGTTTCTCAGAAATTTGTAGGAAGAGCCAT 1546  
Db 1574 TATTAGTCTCAATGAATTTGTTAGCCCGTTTCTCAGAAATTTATAGGAAGAGAGCTAT 1633  
Qy 1547 AAGAATTTAGTGGTATTTCAAGAGACATACAGAAAGAGGGTGAATTTGAGGCTTTGAG 1606  
Db 1634 AAGAATAAGTGGTATTTCAAGAAATCTTATAGAAGAAACCTGAGAACGTTGGAGGCTTTGAG 1693  
Qy 1607 AAATTTGCTATTTGACATATATGAGGGTCAGATTTACTGCTTACTTTGGCCACAGTGGAAC 1666  
Db 1694 AAATTTGCTATTTGACATATATGAGGGTCAGATTTACTGCTTACTTTGGCCACAGTGGAAC 1753  
Qy 1667 AGGAAAGAGTACATTCATCAATATTTCTTTGTGAGCTCTGCCACCTTCTGATGGGTTTGC 1726  
Db 1754 AGGAAAGAGTACATTCATCAATATTTCTTTGTGAGCTGTTGTCACCTCTGATGGGTTTGC 1813  
Qy 1727 ATCTATATATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGAT 1786  
Db 1814 TTCTATATATGGACACAGAGTCTCTGAAATAGATGAATGTTTGAAGCAAGAAATGAT 1873  
Qy 1787 TGGCATTTTGTCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAAGAAATTTATC 1846  
Db 1874 TGGCATATGTCGCCAGTCAATATAAATTTGATGTTTGTGACAGTAGAAGAAATTTATC 1933  
Qy 1847 AATTTTGGCTTCAATCAAGGGATACCGCCCAACATATAATATACAGAGTGCAGAGGT 1906

Db 1934 AATTTTGGCTTCAATCAAGGAATAACAGCCAAACAATATAATTTCAAGAGTGCAGAAAGT 1993  
Qy 1907 TTTTACTAGATTTAGACATGCAGACTATCAAGAGATAACCAAGCTTAAATAATTTAAGTGTGG 1966  
Db 1994 GTTGTGGATCTGCATGCATGCAAGCCATCAAGAGATAATCAAGCGAAAAAGTTAAGCGGTGG 2053  
Qy 1967 TCAAAAAAGAAAGCTGTCTATTAGGAATGCTGTTCTTGGGAACCCAAAGATCTGCTGCT 2026  
Db 2054 TCAGAAAAAGGAAGCTGTGTAGGAATTCAGTTCTCGGGAATCCAAAGATATCTCTGCT 2113  
Qy 2027 AGATGAACCAACAGCTGGAATGACCCCTGTTCTCGACATATTTGTATGGAATCTTTTAA 2086  
Db 2114 AGAGAGCTTACAGCAGGAATGACCCCTGCTCTCGCCATATTTGTTGGAATCTCTTAA 2173  
Qy 2087 ATACAGAAAGCCAAATCGGGTGAAGTGTCTGATCTCTCTCATCTTTCATGATGAAGTGCAT 2146  
Db 2174 GTATAGAAAGGCTTAAACAGAGTGAAGCTGTCTACTCTCTCATGCTGATGAGGCTGACAT 2233  
Qy 2147 TCTTGCAGATAGGAAGCTGTGATATCAAGAGGAATGCTGAAATGTTGTTGTTCTTCAAT 2206  
Db 2234 TCTTGCAGCAGGAAGCTGTCTATATCAAGGAATGCTGAAATGTTGTTGTTCTTCAAT 2293  
Qy 2207 GTTCTCTCAAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTG 2266  
Db 2294 TTTCTTAAAGATAATGGGAAATCGGCTACCGCTGAGCATGTATATAGACAGATACTG 2353  
Qy 2267 TGGCAGAGATCTCTTCTTCTACTGTTTAAACAACATATACCTGGAGTACTTTTATACA 2326  
Db 2354 TGGCAGAGTCACTGCTGCTGCTGTTAGGAGCAGCATACCGCGCAGCGGCTACTGCA 2413  
Qy 2327 ACAGAAATGACCAACAACTTGTATAGTTCCTTTTCAAGGACATGGACAAATTTTCAGG 2386  
Db 2414 GCAGAAATGACAGCAGCTCGTGTACAGCCTGCTTTTCAAGACATGGACAAATTTTCAGG 2473  
Qy 2387 TTTGTTTCTGCTTACAGAGTCAATTTCAATTTGGGTGTCATTTCTTATGCTGTTTCCAT 2446  
Db 2474 CTTGTTTCTGCTTACAGATTCATTTCAACCTGGGTGTTTCTTATGCTGTTTCCAT 2533  
Qy 2447 GAGCAGCTTTGGAAGAGCTATTTTAAAGCTAGAGTTTGAAGCAGAAATTTGACCAAGCAG 2506  
Db 2534 GACAACTTGAAGATGTTATTTTAAAGCTAGAGTTTGAAGCAGAAATTTGACCAAGCAG 2593  
Qy 2507 TTATAGTATTTTACTCAGCAGCCACTGGGAGGAATTTGATTTCAAAATCTTTTATGA 2566  
Db 2594 CTATAGTATTTTACAGCAGCGCGGGAGGAAGAAACAGATTTCAAAATCTTTTATGA 2653  
Qy 2567 AATGGAACAGAGCTTACTTATTTCTGAAACCAAGGCTTCTCTAGTGAAGCAGCCTAG 2626  
Db 2654 AATGGAACAGAGTTTACTTATTTCTCTGAAACGAAAGCTTCTCTAGTGAAGCAGCCTAG 2713  
Qy 2627 CTTTGGAAACAAACAGATGATACATAGCAAGTTTCAATTTCTTTTACCTTGAACAGTGA 2686  
Db 2714 CTTCTGGAAGCAGCAAGTGTCTACGATTTGAAAGTTTCAATTTCTCTCAATTTGAACAG 2773  
Qy 2687 AAGTAAATCAGTGAAGTCTGCTTCTGCTTTTAAATTTTACAGTTTCAAGTTT 2746  
Db 2774 AAGCAATCAGTGGCGCTGTTGCTTCTGCTTTTAAATTTTTCAGTTTTCAGTTT 2833  
Qy 2747 TATGTTTTTGGTTTCACTCTTTTAAATAATGCTGGTTTCCCATCAAACTTGTTCAGA 2806  
Db 2834 TATGTTTTTCTCCATCATTTCTTTTAAATAATGCTGGTTTCCCATCAAACTTGTTCAGA 2893  
Qy 2807 CTTATATTTTCTAAACCTGGAGCAACCAACATTAATACAAACAGTCTGCTTCTTCA 2866  
Db 2894 CTTGTATTTCTAAGGCTGGAGTAACCTCTAATAATACAAACAGGCTGCTGCTTCA 2953  
Qy 2867 AAATTTCTGCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCAGAGCCAGAACATAAT 2926  
Db 2954 AAATTTCTACTGACTCAGATATCAATGGTCTTATTTAGTTTGTGCAACAGGCTGCTTCA 3013  
Qy 2927 GGTGACGATATTAATGACAGTGTATGTTATCGGTGGTCTCCCATAGTGGGCTTTTAA 2986  
Db 3014 GGTGGCAATTTTAAATGACAGTGTATGTTCTGCTCTCTCCTCACAGTGGGCTCTCA 3073





```
/strain="C57BL/6J"
/db_xref="PANTOM DB:B230352G14"
/db_xref="MGI:2411591"
/db_xref="taxon:10090"
/clone="B230352G14"
/sex="male"
/tissue_type="corpora quadrigemina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1..3831
/notes="AFP-BINDING CASSETTE PROTEIN homolog [Homo sapiens]
(SPTR|CAB93535, evidence: FASTV, 89.4%ID, 76.5%length,
match=3750)"
ORIGIN
Query Match 51.3%; Score 2811; DB 11; Length 3831;
Best Local Similarity 86.6%; Pred. No. 0; Mismatches 480; Indels 1; Gaps 1;
Matches 3110; Conservative 0;

QY 107 AGAAAAACATGTCACCTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACCAGAAACACTTCT 166
DB 240 AGAAAAACATGGCTACTGCAATTAGGGATGTGGAGTTTGGAGACAGACCAGAAACACTTCT 299
QY 167 ACTGAAGATTACTTAATTAATGACAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTT 226
DB 300 ACTGAAATTAACCTAAATTAATGAGGACTTAAAGAAAGTAGTGTTCAGGAAATCTTTT 359
QY 227 TCCACTATTTTTTTTATTTGGTTTAAATTAATAGCATGATGCATCCAAATAGAAATA 286
DB 360 TCCTCTATTTTCTTATTTGGCTGATATTAGTAGCATGATGCATCCAAATAGAAATA 419
QY 287 TGAAGAGTGCCATAATAGAACTCAATCTTATGAGCAAGTTTACTCTTTCTAATTAAT 346
DB 420 TGAAGAGGTATCTGATATAGAGTCAAGCCCTATGACAAATTCAGCCCTTTCCAAAGTTAT 479
QY 347 TCTTGGATATCTCCAGTGACTAATATTACAGAGAGCATCATGACAGAAAGTGTCTACTGA 406
DB 480 TCTTGGATACATCCCGTGACTTAACATTACAGAGAGCATTAATGACAGGGTTTCTACGA 539
QY 407 TCATCTACCTGATGTCATAATTTACTGAAGAAATPATCAAAATGAAAGAAATGTTAAACATC 466
DB 540 TCATCTTCCCAAGGTATAGTTTACTCAAGAAATPATCAAAATGAGAAAGAACTGGTAGCCGC 599
QY 467 CAGTCTCTCAAGCCGACCACTTTGTAGGTGTGTTTCAAGACTCCAGTCTCTATGA 526
DB 600 AAGTCTTCTAAGTCCAGCACTTCGTAGGTGTGTTTCAAGACCACTGTCTATGA 659
QY 527 ACTTCGTTTTTTTCTCGATGATGATCCAGTATCTTTCTATTTATATGGAATTCAGAGCTGG 586
DB 660 ACTTCGTTTTTTTCTCGAATGATCCAGTGTCTTTCTATTTATATGAATTCAGAGAGG 719
QY 587 CTGTTCAAAATCATGTGAGGCTGCTCAGTACTGTCCTCAGGTTTCAGAGTTTACAAGC 646
DB 720 CTGTTCAAAAGCATGTGATGCTGCTCAGTACTGTCCTTTGGGGTTTACAGTTCTGCAGGC 779
QY 647 ATCCATAGATGCTGCCATATATACAGTTGAAGACCAATGTTTCTCTTTGGAAGGAGCTGA 706
DB 780 ATCGATAGATGCTGCCATATATACAGTGAAGACCAATGTTTCTGTGTGAGCGAGCTGA 839
QY 707 GTCAATAAGCTGTTTATTCGGGAAACCTGCTGTTGTAGAAATAGATATACCTTTCCCGC 766
DB 840 GTGACCAAGAGCTGATCATGGGAGAGCGCGCTGTTGTGAGATTTGACACCTTCCCGCG 899
QY 767 AGGAGTAATTTTAAATATACCTAGTTATAGCATTTTTCAGCTTTTGGATACTTTTGGCAAT 826
DB 900 AGGGGTATCTCTATCTCTGCTCATAGCCTTCTCGCCCTTCGGCTACTTCTTGGCAAT 959
QY 827 TCATATCGTAGCAAAAAAGAAAAAATAAAGAAATTTTAAAGATAAATAGGACATTC 886
DB 960 CCACATCTGTGGCAGAAAAAGAAAGTTAAAGAAATTTTAAAGATAAATAGGACATTC 1019
QY 887 TGATACCTGCTTTTGGCTTTCCTGGTCTCTCTATATACAGTTTAAATTTTCTTATGTC 946
```

2026 TAGATGAACCAACAGCTGGAATGGACCCCTGTTCTCGACATATTGTATGGAAATCTTTAA 2085  
 Db |  
 2160 TAGACGAGCTTACAGAGAAATGGACCCCTGCTCTCGCCATATTGTTTGGAAATCTTTAA 2219  
 Qy |  
 2086 AATACAGAAAGCCAAATCGGCTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACA 2145  
 Db |  
 2220 AGTATAGAAAGGCTAACAGAGTGAACGGTGTGTTAGTACTCACTTCATGATGAGGCTGACA 2279  
 Qy |  
 2146 TTCTTGCATAGGAAGCTGTGATATCACAGGAATGCTGAATGTCTGTTCTTCAA 2205  
 Db |  
 2280 TTCTTGCCACAGGAAGCTGTATATCACAGGAATGCTGAAGTGTGTTGTTCTTCAA 2339  
 Qy |  
 2206 TGTTCCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACACATGACAAATATT 2265  
 Db |  
 2340 TTTTCTAAAGAGTAAATGGGAAATCGGCTACCGCTGAGCATGTATATGACAGATACT 2399  
 Qy |  
 2266 GTGCCACAGAAATCTTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATAC 2325  
 Db |  
 2400 GTGCCACAGAGTACTGTCTGCTGCTGTTAGGACGACATACCCGCGCTACTGTC 2459  
 Qy |  
 2326 AACAGAAATGACCAACAACTGTGTATAGCTTCCCTTTCAAGGACATGACAAATTTTCAG 2385  
 Db |  
 2460 AGCAGAAATGCCAGAGCTCGTGTACAGCGCTGCTTTCAAAGACATGACAAATTTTCAG 2519  
 Qy |  
 2386 GTTGTGTTTCTGCCCTAGACAGTCAATTAATTTGGGTGTCAATTTCTTATGTTTCCCA 2445  
 Db |  
 2520 GCTTGTGTTTCTGCTCTAGACATTCATTTCAAACCTGGGTGTTATTTCTTATGTTTCCCA 2579  
 Qy |  
 2446 TGACGACTTTTGAAGACGCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAG 2505  
 Db |  
 2580 TGACACATTTGAGAGATGATTTCTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAG 2639  
 Qy |  
 2506 ATTATAGTATTTACTAGAGCCACCTGGAGAGAAATGAGATTCAAAATCTTTTGATG 2565  
 Db |  
 2640 ACTATAGTATTTTACACAGACGCGCGGAGAGAAACAGATTCAAAATCTTTTGATG 2699  
 Qy |  
 2566 AATGAGACAGAGCTTACTTATTTCTTGAACCAAGCTTCTAGTGAAGCAGCATTGA 2625  
 Db |  
 2700 AATGGAACAGAGTTTACTTATTTCTCTGAAACGAAAGCTTCTTCAAGTGAAGCAGCATTGA 2759  
 Qy |  
 2626 GCCTTTTGGAAACACAGATGTATACAAATAGCAAAAGTTTCAATTTCTTACCTTGAACGCTG 2685  
 Db |  
 2760 GCCTCTGGAAGCAGCAAGTGTCTACGATTTGCAAAAGTTTCAATTTCTCTCTCATTTGAACGAG 2819  
 Qy |  
 2686 AAGTAATACAGTGAGTACAGTGTGCTCTGCTTTTAAATTTTTCACAGTTCAGATTT 2745  
 Db |  
 2820 AAAGCAAAATCAGTGCCTGCTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTT 2879  
 Qy |  
 2746 TTAATGTTTTTGGTTTCATCTCTTTTAAATAATGCTGTGTTCCCATCAAACTTTGTTCCAG 2805  
 Db |  
 2880 TTAATGTTTTTCTCCATCATCTTTTAAATAATGCTGTGTTCCCATCAAACTCGTTCCAG 2939  
 Qy |  
 2806 ACTATATTTTCTAAACCTGGAGCAACCAACATATAAATACAAACAAAGTCTGCTTCTC 2865  
 Db |  
 2940 ACTGTATTTCTTAAAGCTGGAGATAAACCCTCATAAATACAAACAAAGCTGCTGCTTC 2999  
 Qy |  
 2866 AAAATTCGTGCTACTCAGATATCAGTGATCTTATAGCTTTTTCACAAAGCCAGCAACATAA 2925  
 Db |  
 3000 AAAATTCGTGCTACTCAGATATCAATGGTCTTATAGTATTTTTCACACACAGCAACATAA 3059  
 Qy |  
 2926 TGTGACGATGATTAATGACAGTGAATATGATATCGGTGCTGCCCATAGTGGCGCTTAA 2985  
 Db |  
 3060 TGTGGCAATGTTTAAATGACAGTGAATATGATGCTGCTGCTCTCTCCTCAGAGTGGCGTCTCA 3119  
 Qy |  
 2986 ATGTGATGCAATTCAGAAAGGATGATGTTTGTGACGCTGTTTTCACAGTACTATGGTTT 3045  
 Db |  
 3120 ATGTGTCGCTTCTGAAAGGATGATGTTTGTGCTGCTGTTTTCACAGTACTATGGTTT 3179  
 Qy |  
 3046 ATTCTTTTACCTATATTAGTGAATATCATTTAGTAACCTACTATCTTTATCATTTTAAATGGA 3105  
 Db |  
 3180 ATTGTTTGGCAGTCATGATGAACATCATTTAGTAACCTACTATCTTTATCATTTTAAATGGA 3239

3106 CTGAACCACTCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCTATAGTTTTTA 3165  
 Db |  
 3240 CTGAACCACTCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCTATAGTTTTTA 3299  
 Qy |  
 3166 AATTGAGCTGATTTTCAAGCAGCTTTGCTTGGAAATCAATTTGTTTCTGCAATGCCACCTT 3225  
 Db |  
 3300 AATTGAGCTATATTTTCAAGCAGCTTTGCTTGGAAATCAATTTGTTTCTGCAATGCCACCTT 3359  
 Qy |  
 3226 ACTTTGCCATGGAATTCAGAGAAATCAATAAGATCAAAAGCTTATATCTCAACTTAAACTTTT 3285  
 Db |  
 3360 ACTTTGCCATGGAATTCAGAGAAATCAATAAGATCAAAAGCTTATATCTCAACTTAAACTTTT 3419  
 Qy |  
 3286 CAGGTCTTTTGCCATCTGCATATATGGAATGGACAGCTGTTGTTGATATCCCTTATTTT 3345  
 Db |  
 3420 CGGGTCTTTTGCCATCAGCTTATTTGGTTTGGCAAGCTGTTGTTGATATTCCTCTGTTTTT 3479  
 Qy |  
 3346 TTATCATTTCTTATTTTGTATGCTAGGAAGCTTATTTGGCAATTTCAATGATGATATATTTT 3405  
 Db |  
 3480 TTGTTGTTGATTTTGTATGCTGGGAAGTTTATTTGCAATTTTCATCATGACATGATTTT 3539  
 Qy |  
 3406 ATACTGTAAAGTTCTCTGCTGCTGTTTTTGGCTTATTTGGTTATGTTGTTTCCATGATTTTC 3465  
 Db |  
 3540 ATCTGTCAAAATTTCTTGTGTTGTTTGTGCTCAATTTGCTTATGTTGCTCTCGTCAATTC 3599  
 Qy |  
 3466 TGTTCATTTATTTGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGT 3525  
 Db |  
 3600 TGTTCAGTACATAGCTTCTGTTTCACTTTTCAAGAAATTTTAAATACCAAGAAATTTTGGT 3659  
 Qy |  
 3526 CATTTATCTATTTCTGTCGAGCGCTTGGCTTGTATTTGCAATCACTGAAATTAACCTTTCTTA 3585  
 Db |  
 3660 CATTTATCTATTTCTGTCGAGCGCTTGGCTTGTGTCGAATCACGGAACCAACTTTCTTTC 3719  
 Qy |  
 3586 TGGATACACAAATTTGCAATCTTCTTCAATGCTGCTTTGTTATCATCTCAATCCAACTATC 3645  
 Db |  
 3720 TGCATATGCAATGAGTACGGCTGCTTTTCAATACACCTTCTGATAGCCATCCAACTATC 3779  
 Qy |  
 3646 CACTTCTAGTGTGCTGATTTCTTTTCAATAAGATTTCTTTGGAAGAAATGATC 3696  
 Db |  
 3780 CTCTCTGGTGTCTGATTTCTTTTCAATAAGGTTCTTTGGAAGAAATATGC 3830

RESULT 3  
 AK034961  
 LOCUS  
 DEFINITION  
 Mus musculus 1529 bp mRNA linear HTC 18-SEP-2003  
 and neck cDNA, RIKEN full-length enriched library, clone:943067009  
 product:ATP-BINDING CASSETTE PROTEIN homolog [Homo sapiens], full  
 insert sequence.  
 AK034961  
 VERSION  
 GI:26330331  
 KEYWORDS  
 HTC; CAP trapper.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
 1  
 Carninci,P. and Hayashizaki,Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE  
 99279253  
 PUBMED  
 10349636

REFERENCE  
 2  
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE  
 20499374  
 PUBMED  
 11042159

REFERENCE  
 3  
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
 Konno,H., Akiyama,Y., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M.,  
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

20530913  
 11076861

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 685-690 (2001)

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

6 (bases 1 to 1529)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Iehii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurahara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Oheato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Ahahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE  
 JOURNAL

Submitted (16-JUL-2001)

Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Please visit our web site for further details.  
 URL: <http://genome.gsc.riken.go.jp/>  
 URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

Location/Qualifiers  
 1..1529  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="PANTOM\_DB:9430067009"  
 /db\_xref="MGI:2399173"  
 /db\_xref="taxon:10090"  
 /clone="9430067009"  
 /tissue\_type="embryonic body between diaphragm region and neck"  
 /clone\_lib="RIKEN full-length enriched mouse cDNA library"  
 /dev\_stage="12 days embryo"

CDS

61..>1527  
 /note="unnamed protein product; ATP-BINDING CASSETTE PROTEIN homolog [Homo sapiens] (SPTP|CAB93535, evidence: PASTY, 89.4%ID, 76.5%length, match=3750)  
 putative"

/codon\_start=1  
 /protein\_id="BAC28896.1"  
 /db\_xref="GI:26330332"  
 /translation="MEOSLILSETKASSVSTMSLWKQVSTIAKPHFLSLKREKSV  
 RAVLLLLIFAVQIFVFFHSPKNAWPIKLVDPDLYFLKPGDKPHKTKYLSLLQNS  
 TDSIDNGLIEFPAHQINWAMFNDSDYVSAAPHSALNVVRSEKDYVFSAVENWVY

CLPVMNIIISNYVHLNVTETAIQTWSTPFIQIBITDIFVKIELYFOAALLGIIVTAMP  
 PYPAMENAHKIKAVTQKLSGLPSAVVGAOVVDIPLFFVWLILMGLSGSPAFPHHG  
 LYFPAKFLANVFCILIANVPSVULFFIYSLASFPEKKILNTEKFWFSYIYNALCAVIT  
 ETTFLOQAVTAVFHTFICILAIPIYPLGLLSIFIKGSKWNKPKNNTINADRLIIVA  
 VIMPYLQCLINWIFLQHYEKIRGHSIRKDPFRALSKAKNNKFFPEPPINDEDESDV  
 KAERLKYKELMGQCCCEKPAIMVNCNLHKEYDDKDFLH"

ORIGIN

Query Match 22.1%; Score 1212.2; DB 11; Length 1529;  
 Best Local Similarity 87.1%; Pred. No. 2.1e-203;  
 Matches 1331; Conservative 0; Mismatches 198; Indels 0; Gaps 0;  
 2508 TATAGTGTATTACTCAGCAGCCATCGAGGAGAAATGGATTCAAAATCTTTTGTGAA 2567  
 Db 1 TATAGTGTATTACACAGCAGCGCGGAGGAGAAACAGATTCAAAATCTTTGTGAA 60  
 2568 ATGGAACAGAGCTTACTTATCTTCTGAAACACAGGCTTCTAGTGAGCACCAGGC 2627  
 Db 61 ATGGAACAGAGTTTACTTATCTTCTGAAACAGAGCTTCTAGTGAGCACCAGGC 120  
 2628 CTTTGGAAAACAACAGATGTATACATAGCAAAAGTTTCATTTCTTACCTTGAACCTGAA 2687  
 Db 121 CTCTGGAAGCAGCAAGTGTCTAGATTGCAAAAGTTTCATTTCTCTCTCATTTGAACCGAGAA 180  
 2688 AGTAAATCATGATGATCAGTGTGCTTCTGCTTTTAATTTTTCACAGTTCAGATTTT 2747  
 Db 181 AGCAATCAGTGCAGCTGTGTGCTTCTGCTTTTAATTTTTCACAGTTCAGATTTT 240  
 2748 ATGTTTTGGTTTCATCAGCTTTTAAATATGCTGGTTCATCAATCAAACTTGTTCAGAC 2807  
 Db 241 ATGTTTTCTCCATCATCTTTTAAATATGCTGGTTCATCAATCAAACTTGTTCAGAC 300  
 2808 TTATATTTTCTAAACCTGGAGACAAACCAATATAAATAAACAAGCTGCTTCTTCAA 2867  
 Db 301 TTATATTTCTAAAGCTGGAGATAACCTATAAATAAACAAGCTGCTTCTTCAA 360  
 2868 AATTCGCTGACTCAGATATCAGTGTATTTAGCTTTTTCACAGCAGCAACAATAAG 2927  
 Db 361 AATTCGCTGACTCAGATATCAATGGTCTTATGAGTTTTCACACAGCAACAATAAG 420  
 2928 GTGACGATTAATGACAGTCACTATGATTCGGTGGCTCCCATAGTCGGCTTTAAAT 2987  
 Db 421 GTGGCAATGTTAAATGACAGTCACTATGATTCGGTGGCTCCCATAGTCGGCTTTAAAT 480  
 2988 GTGATCATTCAGAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTACTGTTTAT 3047  
 Db 481 GTGGCGCTTCGAAAAGGACTATGTTTTCGCTGTTTTCACAGTACTACTGTTTAT 540  
 3048 TCTTTACCTATATTAGTGAATATCATTAAGTAACTACTATCTTTATCAATTAATGTGACT 3107  
 Db 541 TGTTCGCCAGTCATGATGAACATCATTAAGTAACTACTATCTTTATCAATTAATGTGACT 600  
 3108 GAACCATCCAGATCTGGATACCCCATCTTTCAAGAAATTAAGTATAGTTTATAA 3167  
 Db 601 GAAGCATCCAGACCTGGAGTACCCCGTTTCATCAAGAAATTAAGTATAGTTTATAA 660  
 3168 ATTGAGCTGATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTTACTGCAATGCCACTTAC 3227  
 Db 661 ATTGAGCTGATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTTACTGCAATGCCACTTAC 720  
 3228 TTGCGCATGGAATATGACAGAGATCATTAAGTCAAGCTTATCACTCAACTTAACTTCA 3287  
 Db 721 TTGCGCATGGAATATGACAGAGATCATTAAGTCAAGCTTATCACTCAACTTAACTTCA 780  
 3288 GGTCTTTTCCCATCTGTCATATTGGATGGACAAGCTGTTTGTGATATCCCTTATTTT 3347  
 Db 781 GGTCTTTTCCCATCAGCCTATTGGGTGGACAAGCTGTTTGTGATATTTCCCTTGT 840  
 3348 ATCATTTCTATTGATGCTAGGAAGCTTATTTGGCAATTTCAATATGATATATTTTAT 3407  
 Db 841 GTTGTCTGATTTGATGCTGGAAAGTTTATTTGCAATTTCAATGAGCTGATTTTAT 900  
 3408 ACTGTAAGTCTCTGCTGTGGTTTTTTTGGCTTATTTGGTATGTTTCCATCAGTTATTTCTG 3467



Db 949 TCCTTTCCCAATCTTTTAAAGCTGGAAGAGCTAAACATGC-TTTGCCATTGAAGAA 1007  
 QY 4905 TATAGCTTTCTCAAGCAACATGGAACAGAGTTTTGTAGAACTCACTAAAGAACAGAG 4964  
 Db 1008 TATAGC-TTCTCTCAAGCAACATGGAACAGAGTTTTGTAGAACTCACTAAAGAACAGAG 1066  
 QY 4965 GAGGAAGATATAGTTGGACTTTAAACAGCACACTTTGGTGGACGACACACAGAA 5024  
 Db 1067 GAGGAAGATATAGTTGGACTTTAAACAGCACACTTTGGT-GGAACGACACAGAA 1125  
 QY 5025 GATAGAGTAGTATTTGAATTTGTATTTGCTGCTGCTTACTGGGACTTCTTTCTTTT 5084  
 Db 1126 GATAGAGTAGTATTTGAATTTGTATTTGCTGCTGCTTACTGGGACTTCTTTCTTTT 1185  
 QY 5085 CACTTAATTTAACTTTGGTTTAAAGCTTTTATTTGGAATGTTGTAAGTGAACCAAG 5144  
 Db 1186 CACTTAATTTAACTTTGGTTTAAAGCTTTTATTTGGAATGTTGTAAGTGAACCAAG 1245  
 QY 5145 AACGCACCTGGAATTTTCTAAGCTCCCTTAATTTGAATGCTGTGTTGTGTTGTTGCTT 5204  
 Db 1246 AACGCACCTGGAATTTTCTAAGCTCCCTTAATTTGAATGCTGTGTTGTGTTGTTGCTT 1305  
 QY 5205 TTCTTTAAATAAAGCTATGATATTAATTAAGTGAA 5238  
 Db 1306 TTCTTTAAATAAAGCTATGATATTAATTAAGTGAA 1339

## RESULT 5

CD250953 850 bp mRNA linear EST 22-MAY-2003  
 LOCUS AGENCOURT\_14214041 NIH\_MGC.179 Homo sapiens cDNA clone  
 DEFINITION IMAGE:30385589 5', mRNA sequence.

## ACCESSION

CD250953

## VERSION

CD250953.1 GI:31011419

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NIH-MGC <http://mgi.nci.nih.gov/>.

1 (bases 1 to 850)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)

Tissue Procurement: Dr. Michael Brownstein

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: NDAM452 row: 0 column: 06

High quality sequence stop: 719.

Location/Qualifiers

1..850

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30385589"

/tissue\_type="Pituitary"

/lab\_host="DH10B-Ton A ( T1 and T5 phage resistances)"

/clone\_lib="NIH\_MGC.179"

/note="Organ: Brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV

(destroyed); Site 2: NotI; Library is oligo-dT primed and

directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.1 kb. Library was

constructed by (Invitrogen). Note: this is a NIH\_MGC

Library."

## ORIGIN

Query Match 15.3%; Score 837; DB 14; Length 850;  
 Best Local Similarity 99.8%; Pred. No. 2.6e-137;  
 Matches 848; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 4117 ATGGTCTCGCAAAAGCACAAATTAATAA-TTCTGGTTGGTGATATTAAGCAACCACTTCA 4175  
 Db 1 ATGGTCTCGCAAAAGCACAAATTAATAA-TTCTGGTTGGTGATATTAAGCAACCACTTCA 60  
 QY 4176 GGCACGGTATTTTAGGAGATTAATCTTCAGAGACAAGTGAAGATGATGATTCACCTGAAG 4235  
 Db 61 GGCACGGTATTTTAGGAGATTAATCTTCAGAGACAAGTGAAGATGATGATTCACCTGAAG 120  
 QY 4236 TGTATGGGTACTGTCTCAGATAAACCCTTTGTGGCCAGATACACTACATTCAGAGAACT 4295  
 Db 121 TGTATGGGTACTGTCTCAGATAAACCCTTTGTGGCCAGATACACTACATTCAGAGAACT 180  
 QY 4296 TTTGAAATTTATGGAGCTGTCAAAGGAATGAGTGAAGTGAATCAAAGAACTCAATAGT 4355  
 Db 181 TTTGAAATTTATGGAGCTGTCAAAGGAATGAGTGAAGTGAATCAAAGAACTCAATAGT 240  
 QY 4356 CGAATAACACATGCACCTTGATTTAAAGAAACATCTTCAGAGAACTGTAAGAAACTACCT 4415  
 Db 241 CGAATAACACATGCACCTTGATTTAAAGAAACATCTTCAGAGAACTGTAAGAAACTACCT 300  
 QY 4416 GCAGGAATCAAACGAAAGTTGTGTTGCTCTAAGTATGCTAGGGAATCCTCAGATTACT 4475  
 Db 301 GCAGGAATCAAACGAAAGTTGTGTTGCTCTAAGTATGCTAGGGAATCCTCAGATTACT 360  
 QY 4476 TTGCTAGATGAACCACTCTACAGATGATGATCCCAAGCCAAACAGACATGCGGAGCA 4535  
 Db 361 TTGCTAGATGAACCACTCTACAGATGATGATCCCAAGCCAAACAGACATGCGGAGCA 420  
 QY 4536 ATTGGAACCTGATTTAAAGAAAGAGCGGGCTGCTATTCTGACCACCTCAGTATATGGAG 4595  
 Db 421 ATTGGAACCTGATTTAAAGAAAGAGCGGGCTGCTATTCTGACCACCTCAGTATATGGAG 480  
 QY 4596 GAGGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGGCGAGTTAAGATGTATC 4655  
 Db 481 GAGGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGGCGAGTTAAGATGTATC 540  
 QY 4656 GGAACAGTACACACTTAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATATG 4715  
 Db 541 GGAACAGTACACACTTAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATATG 600  
 QY 4716 AAGGACTGGATAGAAAACCTAGAGTAGACCCGCTTCAAAGAGAAATTCAGTATATTTTC 4775  
 Db 601 AAGGACTGGATAGAAAACCTAGAGTAGACCCGCTTCAAAGAGAAATTCAGTATATTTTC 660  
 QY 4776 CCAATGCAAGCGCTCAGGAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGNA 4835  
 Db 661 CCAATGCAAGCGCTCAGGAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGNA 720  
 QY 4836 GATGTTCACTCCCTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATCTTTTGGC 4895  
 Db 721 GATGTTCACTCCCTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATCTTTTGGC 780  
 QY 4896 ATTGAAGATATAGCTTTTCTCAAGCAACATTTGGAAGAGCTTTTGTAGAACTCAATAA 4955  
 Db 781 ATTGAAGATATAGCTTTTCTCAAGCAACATTTGGAAGAGCTTTTGTAGAACTCAATAA 840  
 QY 4956 GAACAGAGG 4965  
 Db 841 GAACAGAGG 850

## RESULT 6

## BC029426

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

BC029426 809 bp mRNA linear HTC 19-NOV-2003  
 Homo sapiens ATP-binding cassette, sub-family A (ABC1), member 5,  
 mRNA (cDNA clone IMAGE:472352), containing frame-shift errors.  
 BC029426  
 BC029426.1 GI:20809520



Tissue Procurement: Dr. James R. Lupeki  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM13589 row: f column: 01  
 High quality sequence stop: 717.  
 Location/Qualifiers  
 1. 872  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6190464"  
 /sex="male"  
 /tissue type="sympathetic trunk"  
 /dev stage="adult, 16 yr"  
 /lab\_host="DH10B"  
 /clone\_lib="Lupski sympathetic trunk"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:  
 NotI; Site 2: SalI; cDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCACGCGTCCG-3' and  
 5'-GACTAGTCTAGATCGGCGGCCCT(15)-3'. Size selected >  
 1 kb for average insert length 1.9 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine); available through Life  
 Technologies."

ORIGIN

Query Match 13.9%; Score 761.4; DB 13; Length 872;  
 Best Local Similarity 97.9%; Pred. No. 5.2e-124;  
 Matches 825; Conservative 0; Mismatches 11; Indels 7; Gaps 5;

QY 4073 TTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGTCCAAATGGTGTGCGCAAAAG 4132  
 DB 1 TTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGTCCAAATGGTGTGCGCAAAAG 60

QY 4133 CACAATTAATATCTTGGTGTGATTAATGAACCACTTACGCCAGGATTTTATGAG 4192  
 DB 61 CACAATTAATATCTTGGTGTGATTAATGAACCACTTACGCCAGGATTTTATGAG 120

QY 4193 AGATTATCTTCAGAGACAAGTGAAGATGATTCATCTGAAGTGTGGTTACTGTCC 4252  
 DB 121 AGATTATCTTCAGAGACAAGTGAAGATGATTCATCTGAAGTGTGGTTACTGTCC 180

QY 4253 TCAGATAAACCCCTTTGTGGCCAGATACATTCAGGAAACATTTTGAATTTATGGAGC 4312  
 DB 181 TCAGATAAACCCCTTTGTGGCCAGATACATTCAGGAAACATTTTGAATTTATGGAGC 240

QY 4313 TGTCAGAGATGATGCAAGTGCATGAAGAAGTCAATAGTGAATACACATGCACT 4372  
 DB 241 TGTCAGAGATGATGCAAGTGCATGAAGAAGTCAATAGTGAATACACATGCACT 300

QY 4373 TGATTTAAAGAACATCTTCAGAGACTGTAAAGAACTACCTCAGGAATCAACGAAA 4432  
 DB 301 TGATTTAAAGAACATCTTCAGAGACTGTAAAGAACTACCTCAGGAATCAACGAAA 360

QY 4433 GTTGTGTTTTGCTTAAGTATGCTAGGAAATCCTCAGATTACTTTGTAGATGAACATC 4492  
 DB 361 GTTGTGTTTTGCTTAAGTATGCTAGGAAATCCTCAGATTACTTTGTAGATGAACATC 420

QY 4493 TACAGGTATGATCCCAAGCCAAACACACATGTGGGAGCAATTCGAATCTGATTTAA 4552  
 DB 421 TACAGGTATGATCCCAAGCCAAACACACATGTGGGAGCAATTCGAATCTGATTTAA 480

QY 4553 AAAACAGAAAGCGGGTGTCTTCTGACCACTCACTATATGAGAGGAGGAGGCTGTCTG 4612  
 DB 481 AAAACAGAAAGCGGGTGTCTTCTGACCACTCACTATATGAGAGGAGGAGGCTGTCTG 540

QY 4613 TGATCGATGATCATCATGGTGTCTGGGCGAGTTAAGATGATCGGAACAGTACACATCT 4672

Db 541 TGATCGATGATCATCATGGTGTCTGGGCGAGTTAAGATGATCGGAACAGTACACATCT 600  
 QY 4673 AAAGAGTAATTTGGAAGAGGCTACTTTTGGAAATTAATTTGAAGGACTGATAGAAA 4732  
 Db 601 AAAGAGTAATTTGGAAGAGGCTACTTTTGGAAATTAATTTGAAGGACTGATAGAAA 660  
 QY 4733 CCTAGAAGTAGACCGCCTTTCAAGAGAAATTCAGTATATTTTCCCAAAATGCAAG-CCGTC 4791  
 Db 661 CCTAGAAGTAGACCGCCTTTCAAGAGAAATTCAGTATATTTTCCCAAAATGCAAGCCGTC 720  
 QY 4792 AGGAAAGTTTTTCTTCTATTTTGGCTTATAAAATTTCTTAAGAGATGTTTCACT-CCCT 4849  
 Db 721 AGGAAAGTTTTTCTTCTATTTTGGCTTATAAAATTTCCCTTAAGAGATGTTTCACTCCCT 780  
 QY 4850 TTCACAATCTTTTTTTTAA--GCTGGAGAGAGCTAAACAT-GCTTTTGCATTGAAGAT 4905  
 Db 781 TTCACAATCTTTTTTTTAAAGAGAGCTAAACATGCTTTTGCATTGAAGAGAA 840  
 QY 4906 ATA 4908  
 Db 841 ATA 843

RESULT 8  
 AY414282  
 LOCUS  
 DEFINITION Homo sapiens ABCA9 gene, VIRTUAL TRANSCRIPT, partial sequence,  
 genomic survey sequence.  
 ACCESSION AY414282  
 VERSION AV414282.1 GI:39770244  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 4620)  
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE 2 (bases 1 to 4620)  
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,  
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,  
 Adams,M.D. and Cargill,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES  
 Location/Qualifiers  
 1. 4620  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>4620  
 /gene="ABCA9"  
 /locus\_tag="HCM5172"

ORIGIN

Query Match 13.8%; Score 754.2; DB 29; Length 4620;  
 Best Local Similarity 50.2%; Pred. No. 5e-123;  
 Matches 2435; Conservative 0; Mismatches 2069; Indels 346; Gaps 19;

QY 150 CAGACCAGAACACTTCTTACTGAAGATTTACTTAATTAATGCAACCAACAAAGAGTAGT 209  
 Db 31 CAAACATGGCTCTTCTCTCTGCAAGAACTGTCTCAAAATATGGAGATGAAGACAGACC 90



QY 210 GTTCAGGAAATCTTTTCCACTATTTTTTTA---TTTTGGTTAAATTAATTAATAGCATG 266  
Db 91 TTGTGGAAATGGCTCTTTTCAATTTCTCTGTACTGTTTCTGTACCTATTTTCTCCAAT 150  
QY 267 ATGCATCCAATAGAAATATGAAGAGTGCCTAATATAGAACTCAATCTATGGACAG 326  
Db 151 TTACATCAAGTTATGACACCTCTCAATGTCCTCAATGATCTGGACGCTGTAGATG 210  
QY 327 TTTACTCTTTCTAATCTAATTTCTGGATATACCTCCAGTGACTAATATTAACAAGCAGATC 386  
Db 211 TTTAATGATACATAATATGTTATGTTGCACTGCACTGAAATCCAAACTACCCAAGAGATA 270  
QY 387 ATGCAGAAAGTGTCTAGTATCATCTACCTGATGTCAATTAATCTAGTAAGATATACAAAT 446  
Db 271 ATGAACAAAGTGGCTTCAGCCCCATCTCTAAAGGAAGAAACAATCATGGGGTGGCTGTAT 330  
QY 447 GAAAAAGAAATGTTAAATCCAGCTCTCTTAAGCCGAGCAACTTTTGTAGGTGGTTTTTC 506  
Db 331 GAAAAAGCATGGATGAATTTGAATTTGAATCTAATCAATAGACGAGTGAGAGTCATCTTT 390  
QY 507 AAAGACTCCATGTCCTATGAACCTTGGTTTTTCTCTG---ATATGATTCAGTATCTTCT 563  
Db 391 ACTGATACCTTCTCTACCAATTTGAAGTTTTCTTTGGGACATAGAATCCCATGATGAAA 450  
QY 564 ATTTATATGGATTCAGAGCTGGCTGTTCAAAATCA-----TGTGAGGCT 608  
Db 451 GAGCACAGACCAATTCAGCTCACTGTCAAGCACTGAATGAATAAATGAAGTGTGAAGT 510  
QY 609 GCTCAGTACTGGCTCAGGTTTCAAGTTTCAAGATCCATAGATGCTGCCATATTA 668  
Db 511 TCAGAGTTCTGGGAGAAAGGCTTTGTAGCTTTTCAAGCTGCCATTAATGCTGTATCAT 570  
QY 669 CAGTTGAAGACCAATGTTTCTTTTGAAGAGCTGGAGTCAACTAAAGCTGTTATTATG 728  
Db 571 GMAATCGCAACAAATCATCTAGTATGGAACAGCTGATGTCACTTCTGTGTACATATG 630  
QY 729 GGAGAAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGTAAATTTTAATATACCTA 788  
Db 631 AAGATATTACCTTTGTTGCCAAGGAGGAGTGTCAACTGAATTTTTCATTTCTTTTTC 690  
QY 789 GTTATAGCATTTTACCTTTTGGATCTTTTGGATCTTTTGGCAATTCATATCGTAGCAGAAAAAGAA 848  
Db 691 ATTTCTTTTCTTACATTTATATATATATATATATATATATATATATATATATATAT 747  
QY 849 AAAAAATAAAGAAATTTTAAAGATATGGGACTTTCATGATACCTGCTTTTGGCTTTCC 908  
Db 748 CAATACATATCTCATTCATGACAATGATGGACTCCGAGAGTCAAGATCTGGCTTTCC 807  
QY 909 TGGGTTCTCTATATACAGTTTAAATTTTCTTATGTCCTTCTTATGGCAGTCAATTCG 968  
Db 808 TGGGTTTGTATGCTGGCTTCACTCTTATCATGTCCTTATATGCTTCTTATTTGTA 867  
QY 969 ACAGCTTCTTTGTTATTTCTCAAAAGTAGCAGCATTTGTGATTTCTGCTTTTTCCTT 1028  
Db 868 AATCTGCACAATTTGCTGCTGACTGTTTGTGATGTTCTTCACTTCTTCTCTC 927  
QY 1029 ATAGATATATCATCTGTAATTTTTCCTTTAATAGTGACACCTCTTTTAAATAATCAAAA 1088  
Db 928 ATGGCTCTGCTCTGATACCTTTAGCTTTTCTGATGAGTGTGTGATAAAGAAACCTTTC 987  
QY 1089 CATGTGGGAATGTTGAATTTTGTGTTACTGTGGCTTTTGTGATTTATTTGGCTTTATGATA 1148  
Db 988 CTATCGGGCTTGGTTGTTTCTCTTATTTGTTCTTTTGGGGATCTCGGATTCCTCAGCA 1047  
QY 1149 ATCTCATAGAAAGTTTCCCAAAATCGTTAGTGGCTTTTCACTCTTTTCTGTCATGTT 1208  
Db 1048 TTGTATACACA---TCTTCTGCTCAATTTTGGAAATGCACTTTGCTCTTCTTAGCCCTTT 1104  
QY 1209 ACTTTTGTGATTTGTTATGTCACAGCTCATGATTTAGAAATTTTAATGAAGTGTCTCA 1268  
Db 1105 GCTTTCACGTGTGGATGGCCAGCTTATACATTTGGACTATGATGGAATCTTAATGCC 1164  
QY 1269 TTTTCAAAATTTGACTGCGAGGCCCATATCTCTTAATTTATACAAATTAATCATGCTCACACTT 1328

Db 1165 CACTTGGATCTTTCACAAATCCATACCTCATATAGTACTCTTTTCATGTTGTTTTT 1224  
QY 1329 AATAGTATATCTATGTCCTCTGTGCTGTCTATCTTGATCAAGTCAITTCAGGGGAATTT 1388  
Db 1225 GACACCTTCTGTATTGTTGATTTGACATTATTTTGCACAAATTTTGGCCGCTGAATAT 1284  
QY 1389 GCTTACCGAATCATCTTTATATTTTCTGAAGCTTTCATATTTGGTCAAGAGCAAGA 1448  
Db 1285 GGACATCATGATGTTCTCCCTTCTTTTCTGAAATCTGTTTGGTTTCAACACGGAAGG 1344  
QY 1449 AATTATGAGGAGTTATCAGAGGGGTAATAATGTAAGAAATATAGTTTGTAGTAAATATT 1508  
Db 1345 GCTAATCATGTGCTTCTTGAGATGAACACAGTTCTGATCTCACTAATGACTGTTTT 1404  
QY 1509 GAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCCATAGAAATTTAGTGTATTTCAGAG 1568  
Db 1405 GAACCAAGTGTCTCCAGAAATTTGTGGGAAAGAGCCATCAGGCCAGGTG----- 1453  
QY 1569 ACATACAGAAAGAGGGTGAATAATGTGGAGGCTTTGAGAAATTTGTCAATTTGACATAT 1628  
Db 1454 -----TGGTCTTTGACATAT 1470  
QY 1629 GAGGTCAGATTAATCTTACTTGGCCACAGTGGAAACAGAAAGAGTACATGATGAT 1688  
Db 1471 GAAGCCAGATCACTGCTCTCTTGGTCACTGAGCTGGAAATTTTGTCCAAATFCCAAC 1530  
QY 1689 ATCTTTTGGGACTCTGCCCTCTGATGGTTTGCATCTATATATGGACACAGATC 1748  
Db 1531 ATACTTGTGGGTTGTGAGTTTCCAACTAGGTTTCACTGCTATTAATCACACACTT 1590  
QY 1749 TCAGAAATAGATGAATGTTTGAAGCAAGAAATATGATGGCATTTGTCACAGTTAGAT 1808  
Db 1591 TCAAGAAATGGCTGATATAGAAATATCAGCAAGTTTCACTGGATTTTGTCCAAATFCCA 1650  
QY 1809 ATACACTTGTGTTTTCAGTACAGCAAAATTTTCAATTTTGGCTTCAATCAAGGG 1868  
Db 1651 GTGCAATTTGGATTTCTCACTGTGAAAGAAACCTCAGGCTGTTTGTCTAAATTAAGAGG 1710  
QY 1869 ATACGAGCCAAACATATATATACAGAAAGTGCAGAGGTTTACTAGATTTTACAGATCGAG 1928  
Db 1711 AT-----TTTGCCACATGAAGTGGAGAA----- 1733  
QY 1929 ACTATCAAGATAACCAAGCTTAAAAATTAAGTGGTGTCAAAAAAGAAAGCTGTCAATTA 1988  
Db 1734 ----- 1733  
QY 1989 GGAAATGCTGTTCTTGGGAAACCAAGATACCTGCTGTAGATGAACCAACAGCTCGAATG 2048  
Db 1734 -----AGAGGTTTGTCTATTGGAATGAACCGACTGCTGGATTG 1770  
QY 2049 GACCCCTGTTTCAGACATATTTGATGGAATCTTTTAAATATACAGAAAGCCAAATCGGGTG 2108  
Db 1771 GATCTCTTTCAAGGCACCGAATATGGAATCTCTCGAAAGAGGGGAAATCAGACAGATGA 1830  
QY 2109 ACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTTGCAGATAGGAAAGCTGTG 2168  
Db 1831 ATCTCTTCAGCACCCAGTTTATAGATGAGGCTGACATTTCTGCGGACAGGAGGTGTTC 1890  
QY 2169 ATATCAAGAGAAATGCTGAAATGTTGTTTCTTCAATGTTTCTCAAAAGTAAATGGGGG 2228  
Db 1891 ATATCCAATGGGAAGCTGAAGTGTGAGGCTCTTCTCTGTTCTTAAAGAAAGAAATGGGGC 1950  
QY 2229 ATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGTGCCACAGAAATCTCTTCTTCA 2288  
Db 1951 ATAGGCTACCATTTAATGTTTGCATCTGATGAAGGTGTGATCCAGAGAGTATAACATCA 2010  
QY 2289 CTGGTTTAAACAACATATACCTGGAGCTACTTTATTATCAACAGAAATGACCAACACTGTG 2348  
Db 2011 CTGGTTTAAAGCAGCACATCTCTGATGCCAAATTTGACAGCAAAAGTGAAGAAATCTTGA 2070  
QY 2349 TATAGCTTGGCTTCAAGGACATGGCAAAATTTTTCAGGTTTGTCTTCTGCCCTAGACAGT 2408

Db 2071 TATATTTTGGCTTTGGAAAGGACAAACAAATTTCCAGNNNNNNNNNNNNNNNNNNNN 2130  
Qy 2409 CATTCAAATTTGGGTGTCAATTTCTTATGGTGTTCCTCATGACGACTTTTGGAGACGTATTT 2468  
Db 2131 TGTTCCTAACCAAGGCATTGAGGATATGTTGTTTCCATAACAACTTTGAATGAGGTGTTT 2190  
Qy 2469 TTAAGCTAGAAGTTGAAGCAGAAATTAACCAAGCAGATATAGTGTATTTACTCAGCAG 2528  
Db 2191 CTGAATTAAGAGAAATCAACTATTGATGAATCAGATATTTGGAATTTGGGCAATTA 2250  
Qy 2529 CCACTGGAGGAAGAAATGGATTCAAAATCTTTTGATGAATGGAACAGAGCTTACTTATT 2588  
Db 2251 CAACCTGATGGGCAAGATATAGAACCTTTGTGACTGGAACAAGTTTGTCTTCC 2310  
Qy 2589 CTTTCTGAAACCAAGGCTTCTCTAGTGAGACCATGAGCCTTTTGGAAACAAACAGATGTAT 2648  
Db 2311 TTCCAGAAACCAAGGAAAC---AATCAGTGGCGTGGCGCTCTGGAGGCAGCAGTCTGT 2367  
Qy 2649 ACAATAGCAAGTTTCATTTCTTACCTTGAACGTGGAAGTAAATCAGTGAATCAGTG 2708  
Db 2368 GCANTAGCAAAAGTTTCGCTTCTTAAAGTTTAAAGAAAGAAAGAAAGCCTGTGGACTATA 2427  
Qy 2709 TTGCTTCTGCTTTAATTTTTTTCAGATTCAGATTTTTATGTTTTGTTTGGTTCATCACTCT 2768  
Db 2428 TTATTGCTTTTGGTATTAGCTTTATCCCTCAACTTTTGGAAACATCTATTCTACGAGTCA 2487  
Qy 2769 TTTAAAATGCTGTGGTCCCATCAACTTGTTCAGACTTATATTTCTTAAACCTGGA 2828  
Db 2488 TATCAGAAAGT---TACCGTGGGAACGTGCTCCAAATACATATCTCTCTCACAGGA 2544  
Qy 2829 GACAAACCAATAAATACAAACCAAGTCTGCTTTCAAATTTCTGCTGACTCAGATATC 2888  
Db 2545 CAACAACCAAGGATCCTCTGACCCATTTACTGGTCAATCAATAAGACAGGGTCAACCAT 2604  
Qy 2889 AGTGATCTTATTAGCTTTTTCACAGCCAGACATATGTTGACGATGATTAATGACAGT 2948  
Db 2605 GATAACTTTTTCATTCOACTGAGGCGACAGACATAGCTATAGAAGTGGATGCCTTTGA 2664  
Qy 2949 GACTATGTATCCGTGCTCCCATAGTGGGGCTTTAAATGTGATGCTCATTCAGAAAAGGAC 3008  
Db 2665 ACTAGAAATGGCACAGATGACCATCTTACATGGTGTCTATCATTTGTGTCAGTGTGAA 2724  
Qy 3009 TATGTTTTTGCAGCTGTTTCAACAGTACTATGTTTTTATCTTTTACCTATATTAGTGAAT 3068  
Db 2725 AAGNN 2784  
Qy 3069 ATCATTAGTAACACTATCTTTATCATTTAAATGTGACTGAAACCATCCAGATCTGGAGT 3128  
Db 2785 NNN 2844  
Qy 3129 ACCCATCTTTCAAGAAATTTACTGATATAGTTTTTAAATTTGAGCTGTATTTTCAAGCA 3188  
Db 2845 NNN 2904  
Qy 3189 GCTTTGCTTGGATCATTTACTGCAATGCCACCTTACTTGGCCATGGAATGCAAG 3248  
Db 2905 TTCTCTGGATACCGATGGCAGCTCTTTCATCTCCATACATTTGCAATGAGCAGATTTGT 2964  
Qy 3249 AATCAATAAGATCAAGCTTATCTCAACTTAACTTTTCAAGTCTTTTGGCATCTGCATAT 3308  
Db 2965 GACTACAAAATAAAGCTCATTTCCAGCTACGGATTTTCAAGGCTCTTACCCTTCTGCATAC 3024  
Qy 3309 TGGATTTGGACAGCTGTTGTGATCCCTTATTTTATCATTTCTTATTTTGTATGTC--- 3366  
Db 3025 TGGTTTGGCAAGCATGGTGGATGTTTCCCTGTACTTTTGTATCTCTCTGCTGTATGCA 3084  
Qy 3367 --TAGGAAGCTTATGGCATTTTCAATTATGGAATATATTTTATCTGTAAAGTCTCTTGC 3424  
Db 3085 ATAATGGATTTATTTTAGCCAGAGGAGATTTATTTATATA-ATTCAAACCTGTATAT 3143  
Qy 3425 TGTGTTTTTGGCTTATTTGGTTATGTTCCATCAGTTTATCTGTTCACTTATTTGCTTC 3484  
Db 3144 TCAATCTCTGTAGTATTGGCTATGCTCATCTCTTCTTCTTCTTGTGACATATGTTGATTC 3203

Qy 3485 TTTCACTTTTAAGAAAATTTTAAATACCAAGAATTTTGGTCAATTTATCTATTCTGTGGC 3544  
Db 3204 ATTCAATTTTTCGCAATGGGAGAAAAAATAGTGGCAATTTGGTCAATTTTCTCT- 3256  
Qy 3545 AGCGTTGGCTTGTATTGCAATCACTGAAATAACTTTCTTTTATGGATACACAATTGCAAC 3604  
Db 3257 -----TAAATGGTCACTTTCTCGATAGTTGTAC 3287  
Qy 3605 TATTCTTTCAATTATGCCCTTTTGTATCATCATTCCAATCTATCCACTTCTAGGTTGCCGTAT 3664  
Db 3288 TGAICTAAATGA----- 3299  
Qy 3665 TTCTTTTCAATAAGNTTTCTTTGGAGATGTACGAAAATAATGTGGACACCTATTAATCCATG 3724  
Db 3300 -----ATATGGAATTTCTAGGGCTATTTTGGACCACGTGTAATACCTCCCTTACAT 3352  
Qy 3725 GGATAGCTTTTCACTAGTCTGTATATC-GCCTTACCTGCACTGTGTACTGTGGATTTTCC 3783  
Db 3353 TGATTTGCTCTCTATTCTATTTTCTGAGCCTTACCTTCTATTCTCTATTCTTTTCA 3412  
Qy 3784 TCTTACAATCTATGAGAAAAAATATGGAGGAGATCAATAAGAAAAGATCCCTTTTCA 3843  
Db 3413 TTCTGCGATCCTAGAAATGAATCTGCAAGGAAGAACTAATGAGAAAGGATCCTGTGTCA 3472  
Qy 3844 GMAACCTTTCAACGAAGTCTTAAAAATAGGAAGCTTCCAGAACCAACCAACAATGAGGATG 3903  
Db 3473 GAATTTCTCAA-----GAAGCAACGCTATTTTCCAAACCCAGAGACCTGNAAGGAG 3526  
Qy 3904 AAGATGAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGCTGATGGGTTGCCAGTGT 3963  
Db 3527 AGGAGGAAGATATCCAGATGGAAGATGAGAACAGTGAATGTCTATGGCTGTGCGAGACT 3586  
Qy 3964 GTGAGGAGAAACCATCATTTATGTCTAGCAATTTTGCATAAAGAAATATGATGAACAAGAAAG 4023  
Db 3587 TTGATGAGACACCCGCTCATTTGCCAGCTGTCTACGGAAGGAATATGAGGCAAAAAGA 3646  
Qy 4024 ATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTTCTGTGA 4083  
Db 3647 AAAATGCTTTTCTTAAAGGAAGAAAATAATGCAACAAGAAATGTCTTTTTGT--- 3702  
Qy 4084 AAAAGAGAGAGACTTTAGGACTATTTGGTCCAAATGTGTGCAAAAGACCAATTTA 4143  
Db 3703 -----GAAGTATAGGACTGTTAGGACACATGAGCTGTTAAAGTACACTATTA 3754  
Qy 4144 ATATTCTGTTGTGTATTTGAACCAACTTTGAGCCAGGCTATTTTATAGGAGATTTCTTT 4203  
Db 3755 AGATGATAA CTGGAGACACAAAACCAACTGACGACAGGTGATTTTGAAGGG----- 3807  
Qy 4204 CAGAGACAAGTGAAGATGATGATCTGAAGTGTATGGGTTACTGCTCCTCAGATAACC 4263  
Db 3808 -----AGCGTGGAGGGGAACCCCTGCGGCTTCTTGGGTTACTGCCCTCAGAGATG 3859  
Qy 4264 CTTTGTGGCAGATPACTACATTTGAGGAACATTTTGAATTTTATGGAGCTGTCAAGGAA 4323  
Db 3860 OGCTGTGCCCAACCTGACAGTGAAGCAGCACCTGGAGGTGTACGCTGCCGTGAAAGTTC 3919  
Qy 4324 TGATGCAAGTGAATGAAGAAAGTATTAAGTGAATAACATGCACTGATTTGATTTAAAG 4383  
Db 3920 TCAGAAAGGGGACGCAATGATCGCCATCAACGTTTGTGGATGCGCTCAAGCTGCAGG 3979  
Qy 4384 AACATCTTCAGAGACTGTAAAGAAACTACTGTCAGGAATCAAAACGAAAGTTGTGTTTG 4443  
Db 3980 ACCAGCTGAAGGCTCCCGTGAAGACCTTTGTCAGAGGGAATAAAGCGAAAGCTGTGCTTG 4039  
Qy 4444 CTCTAAGTATGTAGGAAATCTCAGATTACTTTGTAGATGAACCATCTACAGGTATGG 4503  
Db 4040 TGCTGAGCATCTCTGGGAACCCGCTCAGTGTGTCTCTGATGAGCCCTGACCGGGATGG 4099  
Qy 4504 ATCCCAAGCCAAACACACATGTGGGAGCAATTCGAATCTGCAATTTAAAAACAGAAAGC 4563  
Db 4100 ACCCGAGGGGACGACGAATGTGGCAGGTGATTCGGGCCACCTTTAGAAACACGAGGA 4159

4564 GGCTGCTATTTCTGACCACTCACTATATATGAGAGGAGGAGGAGGCTGTCTGTGATCGAGTAG 4623  
 4160 GGGGCGCCCTCTCTGACCACTACATGCGAGAGGCTGAGCGGTGTGTGACCGAGTGG 4219  
 4624 CTATCATGTGTCTGCGGAGTTAAGATGATCGGAACAGTACCAATCTTAAGAGATAAT 4683  
 4220 CCATCATGTGTGAGGAGGCTGAGATGATTTGGTTCCATCCACACCTGAAACCAAT 4279  
 4684 TTGGAAGAGCTACTTTTTTGGAAATTAATTTGAGAGCTGGATAGAAACCTGAGAGTAG 4743  
 4280 TTGGCAAGACTACCTGCTGGAGATGAAGCTGAAGAACCTGGCAAAA-----TGG 4330  
 4744 ACCGCTTCAAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGGTGAGGAAAGTTTTT 4803  
 4331 AGCCCTCATGAGAGATCCTGAGGCTTTTCCCAAGGCTGCTCAGCAGGAAAGTTCT 4390  
 4804 CTTCATTTTGGCTTATAAAATTCCTAAGAGATGTTTCAAGCAATCTTTT 4863  
 4391 CCTCCCTGATGCTATAAGTTGCTGAGATGTCGAGCTTTTATCAGAGGCTTTCT 4450  
 4864 TTAAGCTGGAAGAGCTAAACATGCTTTTCCCAATGCAAGCGGTGAGGAAAGTTTTT 4923  
 4451 TCAATTAGAGATAGTAAACAGAGTTTGCAGCTGAGAGATGTCGAGCTTCTCAGAGTCTA 4510  
 4924 CATTTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGGAGAT 4973  
 4511 CCCTGGAGCAGGTTTTCTCGAGCTCTCCAGGAGCAGGAGCTGGGTGAT 4560

RESULT 9  
 B0441353  
 LOCUS  
 DEFINITION AGENCOURT\_7838430 NIH\_MGC\_82 Homo sapiens cDNA clone IMAGE:6101304  
 5', mRNA sequence.

ACCESSION B0441353  
 VERSION B0441353.1  
 KEYWORDS 1 GI:21180429  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 773)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@mail.nih.gov  
 Tissue Procurement: CLONTECH  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM2334 row: c column: 01  
 High quality sequence stop: 537.

FEATURES  
 Location/Qualifiers  
 1..773  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6101304"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_82"  
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggcgccctcgcc); Site 2: SfiI (ggccattggcc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor  
 sequence: 5'-CAGCGCATTTAGGCC-3' and 3' adaptor sequence:  
 5'-ATTCTAGAGCGGCGGCGGAGCATG-dt(30)BN-3' (where B = A,  
 C, or G and N = A, C, G, or T). Average insert size  
 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained  
 inserts by PCR. This library was enriched for full-length  
 clones and was constructed by Clontech Laboratories (Palo

Alto, CA)." ORIGIN  
 Query Match 13.5%; Score 741.4; DB 13; Length 773;  
 Best Local Similarity 98.4%; Pred. No. 1.8e-120;  
 Matches 759; Conservative 0; Mismatches 11; Indels 1; Gaps 1;  
 4426 AACGAAGTTGCTGTTTGTCTAAGTATGCTAGGAACTCTCAGATTTACTTCTGCTAGATG 4485  
 3 AACGAAGTTGCTGTTTGTCTAAGTATGCTAGGAACTCTCAGATTTACTTCTGCTAGATG 62  
 4486 AACCATCTACAGTATGATGCCAAAGCCAAACAGACATGTCGCGAGCAATTCGAAGCTG 4545  
 63 AACCATCTACAGTATGATGCCAAAGCCAAACAGACATGTCGCGAGCAATTCGAAGCTG 122  
 4546 CATTAAAAACAGAAAGCGGCTGCTATTCTGACCACTCACTATATGAGGAGGAGGAGG 4605  
 123 CATTAAAAACAGAAAGCGGCTGCTATTCTGACCACTCACTATATGAGGAGGAGGAGG 182  
 4606 CTGCTCTGATCGAGTAGCTATCATGCTGCTGCGGAGTTAAGTATGTCGGAACAGTAC 4665  
 183 CTGCTCTGATCGAGTAGCTATCATGCTGCTGCGGAGTTAAGTATGTCGGAACAGTAC 242  
 4666 AACATCTAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTCGAAGCTGGA 4725  
 243 AACATCTAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTCGAAGCTGGA 302  
 4726 TAGAAAACTTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTCCCAAAATGCAA 4785  
 303 TAGAAAACTTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTCCCAAAATGCAA 362  
 4786 GCCTGAGGAAGTTTCTTCTATTTTGGCTTATAAATTCCTAAGGAGATGTTTCAGT 4845  
 363 GCCTGAGGAAGTTTCTTCTATTTTGGCTTATAAATTCCTAAGGAGATGTTTCAGT 422  
 4846 CCCTTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCATTCGAAGAT 4905  
 423 CCCTTTTCACAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCATTCGAAGAT 482  
 4906 ATAGCTTTTCTAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAGACAGAGG 4965  
 483 ATAGCTTTTCTAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAAGACAGAGG 542  
 4966 AGGAAGATAATAGTTGTGGAACTTTAAACAGCACCTTTTGGTGGGAACGAACAGAG 5025  
 543 AGGAAGATAATAGTTGTGGAACTTTAAACAGCACCTTTTGGTGGGAACGAACAGAG 602  
 5026 ATAGAGTAGTATTTGAATTTGATTTTGGTCTGCTTACTGCGGACTTCTTTCTTTTC 5085  
 603 ATAGAGTAGTATTTGAATTTGATTTTGGTCTGCTTACTGCGGACTTCTTTCTTTTC 662  
 5086 ACTTAATTTAACTTTGCTTTAAAGCTTTTATTTGGAATGTTAATGGAAGACCAAGA 5145  
 663 ACTTAATTTT-ACITTTGTTTAAAGTTTTTTTATTTGGAATGTTAATGGAAGACCAAGA 721  
 5146 ACGCACTTGAATTTTCTAAGCTCCCTAATGGAATGCTGTTGTTGTTG 5196  
 722 ACGCACTTGAATTTTCTAAGCTCCCTAATGGAATGCTGTTGTTGTTG 772  
 RESULT 10  
 B0686259/c  
 LOCUS  
 DEFINITION UI-CF-DUI-adh-m-03-0-UI.s1 UI-CF-DUI Homo sapiens cDNA clone  
 747 bp mRNA linear EST 07-OCT-2002  
 UI-CF-DUI-adh-m-03-0-UI 3', mRNA sequence.  
 ACCESSION B0686259  
 VERSION B0686259.1  
 KEYWORDS GI:23540987  
 SOURCE EST.  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 747)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.reagen.com) or from Open Biosystems  
 (www.openbiosystems.com).  
 Seq primer: M13 FORWARD  
 POLYA=yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..747  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-DUI-adm-m-03-0-UI"  
 /tissue\_type="Primary Lung Epithelial Cells"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-CF-DUI"  
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-CF-DUI is a normalized cDNA library containing the  
 following tissue(s): Primary Lung Epithelial Cells The  
 library was constructed according to Bonaldo, Lennon and  
 Soares, Genome Research, 6:791-806, 1996. First strand  
 cDNA synthesis was primed with an oligo-dT primer  
 containing a Not I site. Double stranded cDNA was ligated  
 to an EcoR I adaptor, digested with Not I, and cloned  
 directionally into pT73-pac vector. The oligonucleotide  
 used to prime the synthesis of first-strand cDNA contains  
 a library tag sequence that is located between the Not I  
 site and the (dT)18 tail. The sequence tag for this  
 library is GCCTGTAGGC.  
 TAG\_TISSUE=Lung Epithelial Cells Tissue nos 359-368  
 TAG\_LIB=UI-CF-DUI  
 TAG\_SEQ=GGCTGTAGGC"

ORIGIN  
 Query Match 13.4%; Score 731; DB 13; Length 747;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-118;  
 Matches 731; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4508 CAAGCCAAACAGCACATGTGCGAGCAATTCGAACCTGATTTAAAAACAGAACGGCGC 4567  
 DB 747 CAAGCCAAACAGCACATGTGCGAGCAATTCGAACCTGATTTAAAAACAGAACGGCGC 688  
 QY 4568 TGTCTTCTGACCACTCATATATGAGGAGGAGGAGGCTGTCTGTATCGAGTAGCTAT 4627  
 DB 687 TGTCTTCTGACCACTCATATATGAGGAGGAGGAGGCTGTCTGTATCGAGTAGCTAT 628  
 QY 4628 CATGTGTCTGGCGAGTTAAGATGTATCGGAACAGTACAACTCTAAAGAGTAATTTGG 4687  
 DB 627 CATGTGTCTGGCGAGTTAAGATGTATCGGAACAGTACAACTCTAAAGAGTAATTTGG 568  
 QY 4688 AAAAGGCTACTTTTGGAAATTAATTAAGGACTGGATAGAAACCTAGAGTAGACCG 4747  
 DB 567 AAAAGGCTACTTTTGGAAATTAATTAAGGACTGGATAGAAACCTAGAGTAGACCG 508  
 QY 4748 CCTTCAAGAGAAATTCAGTATATTTCCCAATCAAGCCGTCAGGAAAGTTTCTTC 4807

Db 507 CCTTCAAGAGAAATTCAGTATATTTCCCAATCAAGCCGTCAGGAAAGTTTCTTC 448  
 QY 4808 TATTTTGGCTTATAAAATTCCTAAGGAAGATGTTAGTCCCTTTTCAAACTCTTTTAA 4867  
 Db 447 TATTTTGGCTTATAAAATTCCTAAGGAAGATGTTAGTCCCTTTTCAAACTCTTTTAA 388  
 QY 4868 GCTGGAAGAGCTAAACATGCTTTTGGCAATTTGAAGAAATATAGCTTTTCTCAAGCAACAT 4927  
 Db 387 GCTGGAAGAGCTAAACATGCTTTTGGCAATTTGAAGAAATATAGCTTTTCTCAAGCAACAT 328  
 QY 4928 GGAACAGGTTTTTGTAGAACTCACTAAAGAACCAAGAGGAGGAGAGATAATAGTTGTGAAC 4987  
 Db 327 GGAACAGGTTTTTGTAGAACTCACTAAAGAACCAAGAGGAGGAGAGATAATAGTTGTGAAC 268  
 QY 4988 TTTAAACAGCACACTTTTGTGGGAAACCAACAGGAAGATAGTAGTAGTATTGGAATTG 5047  
 Db 267 TTTAAACAGCACACTTTTGTGGGAAACCAACAGGAAGATAGTAGTAGTATTGGAATTG 208  
 QY 5048 TATTTGTCGTCGCTTACTTGGGACTTCTTTCTTTTCACTTAATTTTAACTTTGGTTTA 5107  
 Db 207 TATTTGTCGTCGCTTACTTGGGACTTCTTTCTTTTCACTTAATTTTAACTTTGGTTTA 148  
 QY 5108 AAAAGTTTTTTTATTTGAATGTTAACTGGAGAACCAAGAGCGCACTTGAATTTTCTAAG 5167  
 Db 147 AAAAGTTTTTTTATTTGAATGTTAACTGGAGAACCAAGAGCGCACTTGAATTTTCTAAG 88  
 QY 5168 CTCCTTAATTTGAATGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5227  
 Db 87 CTCCTTAATTTGAATGCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 28  
 QY 5228 AATTAAGTGAA 5238  
 Db 27 AATTAAGTGAA 17

RESULT 11  
 BG564439  
 LOCUS 602584385F1 NIH\_MGC\_76 Homo sapiens cDNA clone IMAGE:4712157 5',  
 mRNA sequence.  
 DEFINITION BG564439.1 GI:13572091  
 ACCESSION BG564439  
 VERSION BG564439.1  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 771)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: CLONETECH Laboratories, Inc.  
 cDNA Library Preparation: CLONETECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLCM1553 row: i column: 22  
 High quality sequence stop: 763.  
 Location/Qualifiers  
 1..771  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4712157"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NIH\_MGC\_76"  
 /notes="Organ: liver; Vector: pDNR-LIB (Clontech); Site 1:  
 SfiI (ggccattagcc); Site 2: SfiI (ggccattagcc); 5' and  
 3' adaptors were used in cloning as follows: 5' adaptor





EST.  
 Homo sapiens (human)  
 SOURCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 697)  
 NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 Tumor Gene Index  
 UNPUBLISHED (1997)  
 JOURNAL  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs@email.nih.gov  
 Tissue Procurement: Dr. Steven Gitelis/ Rush Presbiterian, Dept. of  
 Orthopaedics  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 The following repetitive elements were found in this cDNA  
 sequence: 1-25, >AT-richLow\_complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=Yes.

FEATURES  
 source  
 1..697  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-H-EZ1-bca-d-17-0-UI"  
 /tissue\_type="Chondrosarcoma Grade II"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI CGAP Ch2"  
 /note="Organ: Left Pelvis; Vector: pT7T3-Pac (Pharmacia)  
 with a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI CGAP Ch2 is a normalized cDNA library containing the  
 following tissue(s): Chondrosarcoma Grade II. The library  
 was constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT7T3-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 TGATCAGCT.  
 TAG\_TISSUE=grade-2-chondrosarcoma  
 TAG\_LIB=UI-H-EZ1  
 TAG\_SEQ=ATCTAATATG"

ORIGIN  
 Query Match 12.4%; Score 676.4; DB 13; Length 697;  
 Best Local Similarity 99.1%; Pred. No. 5.1e-109;  
 Matches 680; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 Qy 4553 AAACAGAAAGCGGCTCTATTTGACACCTCATATATGAGAGAGCGAGGCTGTCTG 4612  
 Db 697 AAACAGAAAGCGGCGTCTATTTCTGACCACTCATATATGAGAGAGCGAGGCTGTCTG 638  
 Qy 4613 TGATCGAGTAGCTATCATGTGTCTGGCAGTTAAGATGTATCGGACAGTACACATCT 4672  
 Db 637 TGATCGAGTAGCTATCATGTGTCTGGCAGTTAAGATGTATCGGACAGTACACATCT 578  
 Qy 4673 AAACAGTAAATTTGGAAAAGGCTACTTTTTGGAAATTAATTTGAAGACCTGGATAGAAAA 4732  
 Db 577 AAACAGTAAATTTGGAAAAGGCTACTTTGTGGAATTAATTTGAAGACCTGGATAGAAAA 518  
 Qy 4733 CCTAGAGTAGACCGCCTTCAAAGAGAAATTCAGTATATTTTCCCAATCGNAGCCGTCA 4792  
 Db 517 CCTAGAGTAGACCGCCTTCAAAGAGAAATTCAGTATATTTTCCCAATCGAAGCCGTCA 458  
 Qy 4793 GGAAAGTTTTTCTTCTATTTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTC 4852



5'-TCGACCCACGCGTCCG-3' and  
5'-GACTAGTTCGAGCGCGCCCT(15)-3'. Size selected >  
1 kb for average insert length 1.7 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

ORIGIN

Query Match	12.2%;	Score 668;	DB 13;	Length 926;
Best Local Similarity	95.8%;	Pred. No. 1.4e-107;		
Matches 731;	Conservative 0;	Mismatches 21;	Indels 11;	Gaps 4;
QY 3618	GCCTTTGTATCATCATTCATCCATCTTCCACTTCTAGGTGCTGATTTCTTTTCAATAAG	3677		
Db 1	GCCTTTGTATCATCATTCATCCATCTTCCACTTCTAGGTGCTGATTTCTTTTCAATAAG	60		
QY 3678	ATTCTTTGGAAGATGTAGAAATGTGGACACCTATATCCATGGATAGGCTTTCA	3737		
Db 61	ATTCTTTGGAAGATGTAGAAATGTGGACACCTATATCCATGGATAGGCTTTCA	120		
QY 3738	GTAGCTGTTTATATCGCTTTACCTGTCAGTGTGTACTGTGGATTTCTCTTACAATACTAT	3797		
Db 121	GTAGCTGTTTATATCGCTTTACCTGTCAGTGTGTACTGTGGATTTCTCTTACAATACTAT	180		
QY 3798	GAGAAATATGCGGCGACATCAATAGAAAGATCCCTTTTTCAGAACTTTCAACG	3857		
Db 181	GAGAAATATGCGGCGACATCAATAGAAAGATCCCTTTTTCAGAACTTTCAACG	240		
QY 3858	AACTCTAAATATAGAAAGCTTCCAGAACCCAGACCAATGAGATGAAGATGTC	3917		
Db 241	AACTCTAAATATAGAAAGCTTCCAGAACCCAGACCAATGAGATGAAGATGTC	300		
QY 3918	AAAGCTGAAAGCTAAAGCTAAAGCTAAAGCTGATGGTTCGCGATGTTGAGAGAAACCA	3977		
Db 301	AAAGCTGAAAGCTAAAGCTAAAGCTGATGGTTCGCGATGTTGAGAGAAACCA	360		
QY 3978	TCCATTATGTCAGCAATTTGCATAAAGATATGATGACAGAAAGATTTCTTTCTTCA	4037		
Db 361	TCCATTATGTCAGCAATTTGCATAAAGATATGATGACAGAAAGATTTCTTTCTTCA	420		
QY 4038	AGAAAGTAAAGAAAGTGCCAACTAAATACATCTCTTTCTGTGTGAAAAA---GGAGA	4093		
Db 421	AGAAAGTAAAGAAAGTGCCAACTAAATACATCTCTTTCTGTGTGAAAAAAGGCTGGAGA	480		
QY 4094	GATCTTAGGACTATTGGGTCCAAATGGTCTGGGAAAGCACCAATTATATATCTGGT	4153		
Db 481	GATCTTAGGACTATTGGGTCCAAATGGTCTGGGAAAGCACCAATTATATATCTGGT	540		
QY 4154	TGGTGATATTGAACCAACTTCAGGCCAGGTATTTTAGGATTTCTTTCAGAGACAAG	4213		
Db 541	TGGTGATATTGAACCAACTTCAGGCCAGGTATTTTAGGATTTCTTTCAGAGACAAG	600		
QY 4214	TGAAGATGATGATTCACCTAGTGTATGGGTACTGTCTCAGATAAACCCCTTTGTGCC	4273		
Db 601	TGAAGATGATGATTCACCTAGTGTATGGGTACTGTCTCAGATAAACCCCTTTGTGCC	660		
QY 4274	AGATACATCATTTGCAGG---AACATTTGAAATTTATGGAGCTGTCAAGGAA---TGAG	4327		
Db 661	AGATACATCATTTGCAGG---AACATTTGAAATTTATGGAGCTGTCAAGGAAATGAGTG	720		
QY 4328	TGCAAGTGACATGAAGAGTCAAT--AGTCCGAATTAACCATGC	4369		
Db 721	CACAGTGACATGAAGAGTCAATTAATTCGAATTAACCATGC	763		

Search completed: April 12, 2004, 07:51:37  
Job time : 8479 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 12, 2004, 18:02:58 ; Search time 9482 Seconds  
(without alignments)  
7505.723 Million cell updates/sec

Title: US-10-090-458-5

Perfect score: 8426

Sequence: 1 MSTAIREGVWQRTRLTK.....CGTLNSTLWERTQEDRWVF 1642

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgit2.1/USPTO.spool/US10090458/runat 08042004 121552 25470/app query.fasta\_1.1799  
-DB=GenEmbl -QWMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10090458 @CGN 1 1 6133 @runat 08042004 121552 25470 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*

1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sv:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sce:\*  
28: em\_un:\*

29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_nam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8426	100.0	5475	6	AX537473 Sequence
2	8422	100.0	6369	6	AX392931 Sequence
3	8422	100.0	6525	6	AX473847 Sequence
4	8422	100.0	6525	9	AY028897 Homo sapi
5	8421	99.9	5096	9	HSR275973
6	8395	99.6	4929	6	AX417824 Sequence
7	8395	99.6	5262	6	AX417828 Sequence
8	8309	98.6	4917	6	AX537472 Sequence
9	8309	98.6	5463	6	AX537470 Sequence
10	8112	96.3	4785	6	AX417826 Sequence
11	7774	92.3	5170	10	RNO426052
12	7737.5	91.8	4945	10	RNO550165
13	7729	91.7	4929	10	AB097675
14	7674	91.1	5347	10	AF491842
15	6477	76.9	4078	9	HSR806923
16	4930.5	58.5	3347	6	AX833036
17	4930.5	58.5	3347	9	AK094416
18	4847.5	57.5	3268	6	AX714300
19	4847.5	57.5	3268	9	AK056533
20	4663	55.3	2845	9	HSR512612
21	3829	45.4	3112	9	AB067475
22	3481	41.3	5243	9	HSR807592
23	3234	38.4	6181	6	AX473850
24	3229	38.3	5722	6	AX657703
25	3228	38.3	6333	9	AY247065
26	3215.5	38.2	4875	9	AF423307
27	3200.5	38.0	4875	6	AX535879
28	3200.5	38.0	5018	6	AX535877
29	3197.5	37.9	5981	6	AX473849
30	3197.5	37.9	5981	9	AY028899
31	3195	37.9	5335	10	AF498362
32	3192	37.9	5680	6	AX657701
33	3169	37.6	6268	10	AF491299
34	3148	37.4	6141	9	HSR803311
35	3137.5	37.2	1864	9	AK122803
36	3124	37.1	4984	9	AF373250
37	3124	37.1	5211	6	AX575480
38	3124	37.1	5296	6	AX473848
39	3124	37.1	5296	9	AY028898
40	3117	37.0	4854	6	AX538009
41	3117	37.0	5332	6	AX538007
42	3108	36.9	5677	9	AB020629
43	3065	36.4	5149	6	AX392958
44	3044	36.1	5622	6	AX392940
45					

# ALIGNMENTS

RESULT 1



Db 1794 TGTCACAGTTAGATATACATTTTGATGCTTTTGACAGTAGAAGAAATTTTATCAATTTTG 1853  
Qy 581 AlaserileysGlyIleProAlasnasnlelleGlnGluValGlnLysValLeuLeu 600  
Db 1854 GCTTCAATCAAGGGATACACAGCAACAATATAACAAGAGTGCAGAGGTTTTACTA 1913  
Qy 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
Db 1914 GATTAGACATGCAGACTATCAAGATAACCAAGCTAATAAAATTAAGTGGTGCAGAAA 1973  
Qy 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
Db 1974 AGAAGCTGTCTATTAGGAATTTGCTGTTCTTGGGAACCCCAAGATACCTGCTAGATGAA 2033  
Qy 641 ProThrAlaGlyMetAspProCysSerArgHisIleValIleProAsnLeuLeuLysTyrArg 660  
Db 2034 CCAACAGCTGGATGGACCCCTGTTCTCGACATATTGTTATGGAACTCTTTAATAACAGA 2093  
Qy 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 2094 AAAGCAATCGGTGACAGTGTTCAGTACTCATTTCTATGGATGAAGCTGACATTTCTTGCA 2153  
Qy 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2154 GATAGGAAGCTGTGATATCAAGGAATGCTGAATGTGTGGTCTTCAATGTTCCTC 2213  
Qy 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2214 AAAAGTAATCGGGATCGGTACCGCTGAGCATGTACATAGACAAATATTGTGCCACA 2273  
Qy 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
Db 2274 GAATCTCTTTCTTCACTGGTTTAAACAACATATACCTGGAGCTACTTTATTACACAGAA 2333  
Qy 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 2334 GACCAACACTTGTGTATAGCTTGCCTTTCAGACATGACAAATTTTCAGGTTGTTT 2393  
Qy 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2394 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCTATTTCTATGGTGTGTTCCATGACACT 2453  
Qy 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2454 TTGGNAGACGTATTTTAAAGCTAGAGTTGAACAGAAATTTGACCAAGCAGATTATAGT 2513  
Qy 801 ValPheThrGlnGlnProLeuGlnGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2514 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGANTCAAATCTTTTGATGAATGGAA 2573  
Qy 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 2574 CAGAGCTTACTTATTTCTTGAACCCAGGCTTCTCTAGTGACACATGAGCCTTTGG 2633  
Qy 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db 2634 AAACAACAGATGTATACATAAGCAAGTTTCATTTCTTTACCTTGAACACGTGAAGTAAA 2693  
Qy 861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2694 TCAGTGAGATCAGTGTCTCTCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTT 2753  
Qy 881 LeuValHisHisSerPheLysAsnAlaValValProLysLeuValProAspLeuTyr 900  
Db 2754 TTGGTTTCATCTCTTTTAAAAATGCTGTGTGTTCCCATCAACCTTTGTCAGACTTATAT 2813  
Qy 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 2814 TTTCTAAAACCTGGAGACAAACCATATAATACAAAACAAGTCTGCTTCTTCAAAATCT 2873  
Qy 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 2874 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGACATTAATGGTGACG 2933

Qy 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 2934 ATGATTAATGACAGTGACTATGTATCCGGTGGTCCCATAGTGGCGCTTTAATGTGATG 2993  
Qy 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 2994 CATTCAAGAAAAGACTATGTTTTCAGCTGTTTTCAACAGTACTATGGTTATTCTTTA 3053  
Qy 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 3054 CCTATTATTAGTGAATATCATTAGTAACACTATCTTTATCATTTAAATGTGACTGAAC 3113  
Qy 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 3114 ATCCAGATCTGAGTACCCCACTTTCTTCAAGAAATTAAGTATATGTTTTAAAAATTGAG 3173  
Qy 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 3174 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTACTGCAATGCCCTTACTTTGCC 3233  
Qy 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3234 ATGGAAATGACAGAGATCATAGATCAAGCTTATACTCACTTAACTTTTCAGGCTCT 3293  
Qy 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
Db 3294 TTGCCATCTGCATATTGGATTGGACAAGCTGTGTGTATATCCCTTATTTTATCATTT 3353  
Qy 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3354 CTTATTTTGATGCTAGGAAGCTTATGGCAATTCATTTATGATATATTTTATACTGTA 3413  
Qy 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 3414 AAGTTCCTTGTGCTGTTTGTGCTTATGTTATGTTTCCATCAGTATTCTGTTCACT 3473  
Qy 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 3474 TATATTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCATTATC 3533  
Qy 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3534 TATTCGTGGCAGCTTGGCTTGTATTGCAATCACTGAAATAAATCTTCTTTATGGATAC 3593  
Qy 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 3594 ACAATTGCAACTATTCTTCATTTATGCTTTTGTATCATCTTCATTCATCTTCCATCT 3653  
Qy 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
Db 3654 GGTTCCTGATTTCTTTCATAAAGATTTCTTGGAAAGATGTACGAAAAAATGTGACACC 3713  
Qy 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 3714 TATAATCCATGGATAGGCTTTTCAGTAGCTGTATATCGCCTTACTCGCAGTGTACTG 3773  
Qy 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 3774 TGGATTTTCTCTTACATACTATGAGAAAAATATGGAGGCAGATCAATAAGAAAAAGAT 3833  
Qy 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 3834 CCCTTTTTCAGAAACCTTTCAACGAAGTCTAAAAATAGGAAGCTTCCAGAACCCACAGAC 3893  
Qy 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 3894 AATGAGATGAAGATGAAGATGTCAAGCTGAAGACTTAAAGSTCAAGAGCTGATGGT 3953  
Qy 1281 CysGlnCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 3954 TGCCAGTGTGTGAGGAGAAACCATCCATATTATGTTGTCAGCAATTTGCATAAAGATATGAT 4013

QY 1301 AspLysLysAspPheLeuLeuSerArgLysValAlaThrLysTyrIleSer 1320  
DB 4014 GACAAAGAAATTTCTCTTCAAGAAAGTAAAGAAAGTGAACATAATCATCTCT 4073  
QY 1321 PheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
DB 4074 TTCGTGTGAAAAAGAGAGATCTTAGAGATATTGGTCCAAATGGTGTGGCAAAAGC 4133  
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
DB 4134 ACAATTAATAATTTCTGGTGGTGATATTGAACCACTTCAGGCCAGGTATTTTAGGA 4193  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
DB 4194 GATTATTCTTCAGAGACAGTGAAGATGATGATTCACCTGAAGTGTATGGGTACTGCTCT 4253  
QY 1381 GlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
DB 4254 CAGATAAACCCCTTGTGGCCAGATACATCATGTGAGGAACATTTTGAATTTATGGAGCT 4313  
QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
DB 4314 GTCAAGGAATGAGTGCAGTGACATGAAGAGTCAATAGTCGAATAACACATGCATTT 4373  
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
DB 4374 GATTTAAAGAACATCTTCAGAGACTCTAAGAAACTTACCTGCAGGAATCAACAGGAAG 4433  
QY 1441 LeuCysPheAlaLeuSerMetLeuGlyVsnProGlnIleThrLeuLeuAspGluProSer 1460  
DB 4434 TTGTGTTTTGCTTAAGTATGCTAGGGAATCTCAGATTAATTTGCTAGATGAACCATCT 4493  
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
DB 4494 ACAGGTATGGATCCCAAGCAACAGCACATGTGGCGAGCAATTCGAACATGCATTTAAA 4553  
QY 1481 AsnArgLysArgAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCys 1500  
DB 4554 AACAGAAAGCGGGCTGCTATTCTGACCCTCACTATATGAGAGGAGCGAGGCTGTCTGT 4613  
QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
DB 4614 GATCAGTAGCTATCATGGTGTCTGGGAGTTAGATGTATCGGACAGTACACATCTA 4673  
QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
DB 4674 AAGAGTAAATTTGGAAAGGGCTACTTTTGGAAATTTAAATTTGAAGGACTGGATAGAAAC 4733  
QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
DB 4734 CTAGAAGTAGCCGCTTCAAGAGAAATTCAGTATATTTCCCAATGCAAGCCGTGAG 4793  
QY 1561 GluSerPheSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
DB 4794 GAAAGTTTCTCTATTTGGCTTATAAAATTTCTAAGGAAGATGTTTCAGTCCCTTTCA 4853  
QY 1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600  
DB 4854 CAATCTTTTTAAGCTGAAGAACTAAACATGCTTTTGGCATTGAAGAAATATAGCTTT 4913  
QY 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
DB 4914 TCTCAAGCAACATTGGAACAGGTTTTTGTAGAACTCACTAAAGAACCAAGAGGAGGAAGAT 4973  
QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyrTrpGluArgThrGlnGluAspArgVal 1640  
DB 4974 AATAGTTGTGGAACTTTAAACAGCACACTTTGGTGGGAACGAACACAGAAGATAGAGTA 5033  
QY 1641 ValPhe 1642  
DB 5034 GTATT 5039

RESULT 2

AX392931  
LOCUS AX392931 6369 bp DNA linear PAT 23-MAR-2002  
DEFINITION Sequence 33 from Patent WO0212340.  
ACCESSION AX392931  
VERSION AX392931.1 GI:19700978  
KEYWORDS Homo sapiens (human)  
SOURCE  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1  
AUTHORS Yue, H., Thornton, M., Ramkumar, J., Tang, Y. T., Azimzai, Y.,  
Baughn, M. R., Yang, J., Yao, M. G., Lal, P., Wallia, N. K., Gandhi, A. R.,  
Hafalia, A. J., Nguyen, D. B., Patterson, C., Elliott, V. S.,  
Tribouley, C. M., Lu, D. A., Xu, Y., Reddy, R., Hernandez, R.,  
Borowsky, M. I., Lo, T. P., Lu, Y., Policky, J. L., Greene, B. D.,  
Sanjanwala, M. S., Rammann, B. E., Burford, N., Ison, C. H., Lee, E. A.,  
Ding, L., Das, D., Kallik, D. A., Khan, F. A. and Seilhamer, J. J.  
JOURNAL Patent: WO 0212340-A 33 14-FEB-2002;  
FEATURES  
Location/Qualifiers  
1..6369  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 2798241CB1"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 6369  
Score: 8422.00 Matches: 1641  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 99.94% Mismatches: 0  
Query Match: 99.95% Indels: 0  
DB: 6 Gaps: 0  
US-10-090-458-5 (1-1642) x AX392931 (1-6369)  
QY 1 MetSerThrAlaIleArgGluValGlyValTtoArgGlnThrArgThrIleuLeuLys 20  
DB 1238 ATGTCCACTGCAATTAGGGAGTAGAGTTGGAGACAGACCAGAACATCTTACTGAG 1297  
QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu 40  
DB 1298 AATTACTTAATTAATGAGAACCAAAAGAGTAGTGTTCAGAAATCTTTTCCACTA 1357  
QY 41 PhePheLeuPheTrpIleLeuIleLeuSerMetMetHisProAsnLysLysTyrGluGlu 60  
DB 1358 TTTTCTTTTATTTTGGTTATATTAATTAATGATGATGATGATGATGATGATGAT 1417  
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
DB 1418 GTGCCCTATATAGAACTCAATCTATGGCAAGTTTACTCTTTCTTAATCTTCTTGA 1477  
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
DB 1478 TATCTCCAGTGACTAATATTAAAGCAGCATCATGCAAGAAAGTGTCTACTGATCATCTA 1537  
QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
DB 1538 CCTGATGTCATTAATTAATGAGAAATATCAATGAAAGAAATGTTAAACATCCAGTCTC 1597  
QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
DB 1598 TCTAAGCCGAGCAACTTTGTAGGTGTGTGTTCAAAGAGATCCCATGCTCTTATGAACCTCT 1657  
QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
DB 1658 TTTTCTTCCGATGATGATCCAGTATCTTCTATTTATGGAATCAAGAGCTGGCTGTCTCA 1717  
QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
DB 1718 AAATCATGTGAGGTGCTCAGTACTGCTCCTCAGGTTTCACAGTTTACAGAGTCCCATTA 1777

QY 181 AspAlaAlaIleIleGlnIleuThrAsnValSerLeuTyrIleValGluSerThr 200  
DB 1778 GATGCTGCCATTATACAGTTGAAGCAATGTTCTTTGGAAGAGCTGGAGTCAACT 1837  
QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrProArgGlyVal 220  
DB 1838 AAGCTGTTATTATGGAGAAACTGCTGTTGTAGAAATAGATACCTTTCCCGGAGAGTA 1897  
QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
DB 1898 ATTTTAATATACCTAGTTATAGCAATTTTCACTTTTGGATACCTTTTGGCAATTCATATC 1957  
QY 241 ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
DB 1958 GTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATATATGGACTTCATGATCT 2017  
QY 261 AlaPheTyrLeuSerTyrValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
DB 2018 GCCTTTTGGCTTTCCCTGGGTCTCTATATACAGTTTAAATTTTCTTATGTGCCCTCTT 2077  
QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
DB 2078 ATGGCAGTCATTGGCAGACTCTTTGTTATTTCCTCAAGTAGCAGCAATGTGTATATT 2137  
QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
DB 2138 CTGCTTTTTCCTTTATGATATATCATCTGTATTTTTCCTCAAGTAGCAGCAATGTGTATATT 2197  
QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
DB 2198 TTTTAAATAATCAAAACATGTGGGATAGTTGAATTTTGTGTACTGTGGCTTTTGATTT 2257  
QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTyrLeuPheSer 360  
DB 2258 ATTGGCTTATGATAATCCTCATAGAAAGTTTTCCTCAAGTAGCAGCAATGTGTATATT 2317  
QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
DB 2318 CCTTTCTGTCTACTTTTGTGATTTGATTTGATTTGACAGCTCATGCAATTTAGAAATTT 2377  
QY 381 AsnGluGlyValAserPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
DB 2378 AATGAAGGTCTCTCAATTTCAAAATTTGCTGCAGGCCCATATCTCTAATTTATCAATT 2437  
QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
DB 2438 ATCATGCTCACATTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTGATCAAGTC 2497  
QY 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTyr 440  
DB 2498 ATTTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCATATTGG 2557  
QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
DB 2558 TCAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGCAATGTTAATGGAATATTAGT 2617  
QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
DB 2618 TTTTAGTGAATATTATGAGCAGTTTCTTCAGAAATTTGTAGGAAGAAAGACCAATAGAAAT 2677  
QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
DB 2678 AGTGGTATTTCAGAAACATACAGAAAGAGGGGTGAAATGTGGAGCTTTGAGAAATTTG 2737  
QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
DB 2738 TCATTTGACATATATGAGGGTCAGATTACTGCTTTACTTGGCCACAGTGGAAACAGAAAG 2797  
QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
DB 2798 AGTACATTGATGAATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 2857  
QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560

DB 2858 TATGGCACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATTGGCAAT 2917  
QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
DB 2918 TGTCCACAGTTAGATATACACTTTGATGTTTTCACAGTAGAAGAAAAATTTATCAATTTTG 2977  
QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
DB 2978 GCTTCAATCAAGAGATACCAGCCACATATATATCAAGAGTGCAGAGGTTTTACTA 3037  
QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys 620  
DB 3038 GATTTAGACATGCAGACTATCAAGATAACCAAGCTAAAAAATTAAGTGTGTGTCAAAAA 3097  
QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuAspGlu 640  
DB 3098 AGAAAGCTGTCAATTAGGAATTTGCTTTTGGGAACCCAAAGATACCTGCTGCTAGATGAA 3157  
QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLeuLysTyrArg 660  
DB 3158 CCAACAGCTGAATGGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAAAATACAGA 3217  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
DB 3218 AAAGCCAAATCGGGTGCAGAGTGTTCAGTACTCATTTTCATGGATGAAGCTGATTTCTGCA 3277  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
DB 3278 GATAGGAAGCTGTGATATCACAGGATGCTGAAATGTGTGTTCTTCAATGTTCTC 3337  
QY 701 LysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 3338 AAAAGTAAATGGGGATTCGGCTACCGCTGAGATGTACATAGACAAATATTGTGCCACA 3397  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
DB 3398 GAATCTCTTCTCTACCTGGTTAAACACATATACCTGGAGTACTTTATTCACACAGAT 3457  
QY 741 AspGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
DB 3458 GACCAACAATTTGTATAGTCTGCCCTTCAAGGACATGACAAATTTTCAGGTTTCTTT 3517  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 3518 TCTCCCTTAGACAGTCAATCAATTTGGGTGTCATTTCTTATGTTGTTCCATGACGACT 3577  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
DB 3578 TTGGAAGACGTATTTTAAAGCTAGAAGTTGAAGCAGAAATTCACCAAGCAGATTATAGT 3637  
QY 801 ValPheThrGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 3638 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGAATCAAAATCTTTTGTAGAAATCGAA 3697  
QY 821 GlnSerLeuLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
DB 3698 CAGAGCTTACTTATTTCTGAAACCAAGGCTCTCTAGTGAGCACCATGAGCCTTTGG 3757  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
DB 3758 AAACAACAGATGTATACATAGCAAGTTCATTTCTTACCTTGAACACCTGGAAGTAAA 3817  
QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 3818 TCACTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAGTTCAGATTTTATGTTT 3877  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
DB 3878 TTGGTTTCACTCTTTTAAAAATGTCTGTGTTCCCATCAAACTTGTTCAGACTTATAT 3937  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920

3938 TTTCTAAACCTGGACACCAACACATATAACAAACCAAGCTCGTCTCTTCAAAATTC 3997  
921 AlaAspSerAspIleSerAspLeuIleSerPheThrSerGlnAsnIleMetValThr 940  
3998 GCTGACTCAGATCATGATCTTATTAGCTTTTTCACAAAGCCAGAACATATGCTGACG 4057  
941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
4058 ATGATTAATGACAGTGACTATGATCCGCTGGCTCCCATAGTCCGGCTTTAAATGCTGTG 4117  
961 HisSerGluIysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
4118 CATTCAGAAAGAGACTATGTTTTGCGAGCTGTTTTCAACAGTACTATGCTTTATCTTTA 4177  
981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
4178 CCTATTATTAGTGNATCATATTAGTACTACTATTCTTATCATTTTAAATGTGACTGAAACC 4237  
1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheIysIleGlu 1020  
4238 ATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTTACTGATATAGTTTTTAAATTCGAG 4297  
1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
4298 CTGTATTTTCAACGACCTTTGCTGGAAATCATTTGTTACGCAATGCCACCTTACTTTGCC 4357  
1041 MetGluAsnAlaGluAsnHisIysIleIysAlaTyrThrGlnLeuLeuSerGlyLeu 1060  
4358 ATGGAAATGCGAGAAATCATTAAGATCAAAAGCTTATCTCAACTTAAACTTTTCAGTCTT 4417  
1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProIlePhePheIleIle 1080  
4418 TTGCCATCTGCATTTGGATTTGGCAAGCTGTTGTTGATATCCCTTATTTTATCATTT 4477  
1081 LeuIleLeuMetLeuGlySerLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
4478 CTATTATTTGATCTAGGAAGCTATTGGCATTTTCAATTTGATTTATTTTATATCTGTA 4537  
1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
4538 AAGTTCCTGCTGCTGTTTTTGGCTTATTTGGTTATGTTTCCATCATGTTATCTGTTCAT 4597  
1121 TyrIleAlaSerPheThrPheIysIysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
4598 TATATTGCTTTCTTCCCTTTTANGAATAATTTAAATACCAAGAAATTTTGGTCATTTATC 4657  
1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
4658 TATTCGTGGCAGCGTTGGCTTGTATTGCAATCACTGMAATAACTTTCTTTATGGGATAC 4717  
1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
4718 ACAATTTGCAACTATCTTCAATATGCTTTTATGCTTATCATCTTCCAAATCATTCACCTTCTA 4777  
1181 GlyCysLeuIleSerPheIleIysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
4778 GGTTCCTGATTTCTTTTCATAAGATTTCTTGAAGAATGTTACGAAATAATGTGACACACC 4837  
1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
4838 TATAATCCATGGGATAGGCTTTTCAGTAGCTGTTTATATCGCTTACTCGCAGTGTGACTTG 4897  
1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
4898 TGGATTTTCTCTTACAACTATATGAGAAAAATAATGAGGCGAGATCAATAAGAAAAAGAT 4957  
1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
4958 CCCTTTTTCAGAAACCTTTCAACGAAAGTCTTAAAAATAGGAAGCTTCCAGAACCCACAGAC 5017  
1261 AsnGluAspGluAspGluAspValLysAlaGluArgLysValLysGluLeuMetGly 1280  
5018 AATGAGGATGAAGATGAGATGTCAAAGCTGAAAGACTTAAAGGTCAAGAGCTGATGGGT 5077

1281 CysGlnCysCysGluGluIysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
5078 TGCAGGTGTTGTGAGGAGAAACCATTCATTTATGGTCAGCAATTTGCATAAAGATATGAT 5137  
1301 AspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer 1320  
5138 GACAAAGAAAGATTTCTTCTTTCAAGAAAGTAAAGAAAGTGCAACTAAATACATCTCT 5197  
1321 PheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
5198 TTCTGTGTGAAAGAGAGAGATCTTAGGACTATTGGTCCAAATGGTGTGCTGCAAAAGC 5257  
1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
5258 ACAATATTATAATTCTGTTGGTGATATTGAACCAACTTCAGGCCAGGTATTTTAGGA 5317  
1361 AspTyrSerSerGlnThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
5318 GATTATTCTTCGAGACAAAGTGAAGATGATGATCTACTGAAGGTATGGGTACTGCTCCT 5377  
1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
5378 CAGATAAACCTTTGTGGCCAGATCTACATTCGACAGAACATTTTGAATTTATGGAGCT 5437  
1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
5438 GTCAAGAAATGAGTGCAGTGCATGAAAGAAAGTCAATAGTCAATAACACATGCACCT 5497  
1421 AspLeuLysGlnHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
5498 GATTTAAAGAACATCTTTCAGAAAGCTGTAAAGAAACTACTGACGAGATCAACACGAAAG 5557  
1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
5558 TTGTGTTTTCCTCTAAAGTATGCTAGGAAATCCCTCAGATTTACTTTGTAGTAGAACCATCT 5617  
1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
5618 ACAGGTATGATGCCAAGCCAAACAGCACATGTGGCGCAATTCGAACTGCATTTAAA 5677  
1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500  
5678 AACAGAAAGCGGCTGTAATTCCTGACCATCTATATGAGGAGGAGGAGGCTGTCTGT 5737  
1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
5738 GATCGAGTAGCTATCATGGGTCTCTGGCAGTTAAGATGATCGGAACAGTACAACTCTA 5797  
1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
5798 AAGAGTAAATTTGAAAAAGGCTACTTTTGGAAATTAATTTGAAGACTCGATAGAAAC 5857  
1541 LeuGluValAspArgLeuGlnArgGluIleGluTyrIlePheProAsnAlaSerArgGln 1560  
5858 CTAGAAGTAGACCCCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAACCGCTCAG 5917  
1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
5918 GAAAGTTTTTCTTCTATTGCTTATAAAATTTCTTAAGGAAGATGTTTCAGTCCCTTCA 5977  
1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600  
5978 CAATCTTTTTTAAAGCTGGAAGAGCTTAAACATGCTTTTCCCATTTGAAGATATAGCTTT 6037  
1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
6038 TCTCAACAACATTTGGAACAGGTTTTTGTAGAACTCCTTAAAGAACAGAGAGGAGAGAT 6097  
1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640  
6098 AATAGTTGTGAACTTTTAAACAGCACACTTTGTGGGAACGAAACAAAGAAAGATAGAGTA 6157



```

QY 1641 ValPhe 1642
Db 6158 GRATIT 6163

RESULT 3
AX473847 LOCUS AX473847 6525 bp DNA linear PAT 09-AUG-2002
DEFINITION Sequence 1 from Patent WO0246458.
ACCESSION AX473847
VERSION AX473847.1 GI:22208006
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Densfle, P., Rosier-Montus, M.F., Prades, C., Arnould-Reguigne, I.,
Duverger, N., Allikmets, R. and Dean, M.
TITLE Nucleic acids of the human abca5, abca6, abca9, and abca10 genes,
vectors containing such nucleic acids and uses thereof
JOURNAL Patent: WO 0246458-A 1 13-JUN-2002;
Aventis Pharma S.A. (FR) ; The Secretary, Department of Health and
Human Services. (US)
FEATURES
source
location/Qualifiers
1..6525
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 6525
Score: 8422.00 Matches: 1641
Percent Similarity: 99.94% Conservative: 0
Best Local Similarity: 99.94% Mismatches: 1
Query Match: 99.95% Indels: 0
DB: 6 Gaps: 0

US-10-090-458-5 (1-1642) x AX473847 (1-6525)

QY 1 MetSerThrAlaileArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20
Db 1011 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGACACAGACAGACCAACTTCTACTGAAG 1070

QY 21 AsnTyrLeuileLysCysArgThrLysLysSerSerValGlnGlnLeuLeuPheProLeu 40
Db 1071 AATTACTTAATTAATGTCAGAACCAACCAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 1130

QY 41 PhePheLeuPheTrpLeuileLeuileSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 1131 TTTTITTTTATTTGGTTTATATATATTAGCATGATGCATCCAAATAGAAATATAGAGAA 1190

QY 61 ValProAsnileGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuileLeuGly 80
Db 1191 GTGCCTAATATAGAACTCAATCCTATGACCAAGTTACTCTTCTTAATCTAATCTTGA 1250

QY 81 TyrThrProValThrAsnileThrSerSerileMetGlnLysValSerThrAspHisLeu 100
Db 1251 TATACTCCAGTGACATAATATACAGCAGCATCATGCAAGAGTGTCTACTGATCATCTA 1310

QY 101 ProAspValileileThrGluGluTyrTrpAsnGluLysGluMetLeuThrSerSerLeu 120
Db 1311 CCTGATGTCATAATTACTGAAGAAATATACAAATGAAAAAGAAATGTTAAACATCCAGTCTC 1370

QY 121 SerLysProSerAsnPheValGlyValPheLysAspSerMetSerTyrGluLeuArg 140
Db 1371 TCTAAGCCGAGCAACTTTGTAGTGTGTTTCAAGAGACTCCATGCTCTATGAATCTCGT 1430

QY 141 PhePheProAspMetileProValSerSerileTyrMetAspSerArgAlaGlyCysSer 160
Db 1431 TTTTTCCTGATGATGATCCAGTACTCTCTATTATATATGGAATCAAGAGCTGGCTGTCA 1490

QY 161 LysSerCysGluAlaalaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerile 180

```

```

Db 1491 AAATCATGTGAGGCTGCTCAGTACTGCTCAGGTTTACAGTTTACAGCATCCATA 1550
QY 181 AspAlaAlaileileGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200
Db 1551 GATGCTGCCATTATACAGTTGAAGCAATGTTCTCTTTTGAAGAGAGTGGAGTCAACT 1610
QY 201 LysAlaValileMetGlyGluThrAlaValValGluileAspThrPheProArgGlyVal 220
Db 1611 AAAGCTGTTATATGAGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGATA 1670
QY 221 IleLeuileTyrLeuValileAlaPheSerProPheGlyTyrPheLeuAlaileHisile 240
Db 1671 ATTTTAATATACCTAGTTATAGCATTTTCCCTTTTGGTACTTTTGGCAATTCATATC 1730
QY 241 ValAlaGluLysGluLysLysileLysGluPheLeuLysileMetGlyLeuHisAspThr 260
Db 1731 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATATGAGGACTTCAATGACT 1790
QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrTrpSerLeuilePheLeuMetSerLeuLeu 280
Db 1791 GCCTTTTGGCTTTTCCCTGGGTCTCTCTATATACAAAGTTTAAATTTTCTATGCTCCCTCT 1850
QY 281 MetAlaValileAlaThrAlaSerLeuLeuPheProGlnSerSerSerileValilePhe 300
Db 1851 ATGGCAGTCATTCGCAGACTTCTTTGTTATTTTCTCAAAGTAGCAGCAATTGTGATTT 1910
QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320
Db 1911 CTGCTTTTTCCTTTTATGGAATATCATCTGTATTTTTCCTTAATGCTGACACCTCTT 1970
QY 321 PheLysLysSerLysHisValGlyileValGluPhePheValThrValAlaPheGlyPhe 340
Db 1971 TTTTAAAAAATCAAAACATGTGGGAATAGTGAATTTTGTACTGTGGCTTTTGATTT 2030
QY 341 IleGlyLeuMetileileLeuileGluSerPheProLysSerLeuValTrpLeuPheSer 360
Db 2031 ATTGGCTTATGATAATCTCTATAGAAAGTTTCCAAATCGTAGTGTGGCTTTTCAGT 2090
QY 361 ProPheCysHisCysThrPheValileGlyileAlaGlnValMetHisLeuLeuAspPhe 380
Db 2091 CCTTCTGTCTGCTGCTACTTTTGTGATGTTATTTGCACAGGTCATGCAATTTAGAAGATTT 2150
QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuLeuileThrile 400
Db 2151 AATGAAGGTGCTTCATTTTCAAAATTTGACTGCAGGCCCATTCCTCTAATATTACAAAT 2210
QY 401 IleMetLeuThrLeuAsnSerilePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420
Db 2211 ATCATGCTCACACTTAATAGTATATCTATGCTCTCTGCTGCTATCTTGATCAAGTC 2270
QY 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440
Db 2271 ATTCCAGGGGAAATTTGGCTTACGGAGATCATCTTATATTTTCTGAAGCCTTCATATGG 2330
QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnileSer 460
Db 2331 TCAAGAGTAAAGAAATATATGAGGAGTTATCAGAGGCAATGTTAATGGAATATTAGT 2390
QY 461 PheSerGluileileGluProValSerSerGluPheValGlyLysGluAlaileArgGly 480
Db 2391 TTTTAGTGAATATTATGAGCCAGTTTCTTCTCAGAAATTTGTAGGAAAAAGACCATAGAAAT 2450
QY 481 SerGlyileGlnLysThrTyrArgLysLysGlyValAsnValGluAlaLeuArgAsnLeu 500
Db 2451 AGTGGTATTTCAGAGACATACAGAAAGAGGGGTGAATATGAGGCTTTGAGAAATTTG 2510
QY 501 SerPheAspLysThrGluGlyGlnileThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520
Db 2511 TCATTTGACATATATGAGGCTCAGATTACTGCTTACTTGGCCACAGTGGACAGGAAAG 2570
QY 521 SerThrLeuMetAsnileLeuCysGlyLeuCysProProSerAspGlyPheAlaSerile 540

```

2571 AGTACATTGAAATATTTCTTTGTGGACTCTGCCCACTTCTGTATGGGTTTGGCACTATA 2630  
541 TTTGlyHisATgValserGluileAspGluMetPheGluAlaAArgLysMetIleGlyile 560  
2631 TATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAATGATGGCAAT 2690  
561 CysProGlnLeuAspIleHisPheAspValleuthrValGluGluAAsnLeuSerIleleu 580  
2691 TGTCACACAGTTAGATATACATCTTTGATGTTTGTGACAGTAGAAGAAAATTTATCAATTTTG 2750  
581 AlaSerIleLysGlyIleProAlaAsnAsnIleleGlnGluValGlnLysValleuLeu 600  
2751 GCTTCAATCAAGGATACCCCAACAAATATATACAAAGAGTGCAGAGGTTTACTA 2810  
601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnIys 620  
2811 GATTTAGACATGCAGACTATCAAGATTAACCAAGCTAAATAATTAAGTGGTGGTCAAAA 2870  
621 ArgLysLeuSerLeuGlyIleAlaValleuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
2871 AGAAAGCTGTCTATTAGGAATTCCTGTTCTTGGGAACCCCAAGATACCTGCTAGATGAA 2930  
641 ProThrAlaGlyMetAspProCysSerArgHisIleValtrAsnLeuLeuLysIleArg 660  
2931 CCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGAATCTTTTAAATACAGA 2990  
661 LysAlaAsnAtqValThrValPheSerThrHisPheMetAspGluAlaAspIleleuAla 680  
2991 AAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTGCA 3050  
681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
3051 GATAGAAAGCTGTGATATCAACAGGAATGCTGAAATGTTGGTTCCTCAATGTTCTCTC 3110  
701 LysSerLysTrpGlyIleGlyTyArgLeuSerMetTyIleAspLysTyCysAlaThr 720  
3111 AAAAGPAAATGGGGATCGCTACCGCTGACATGACATGATAGACAAATATTTGGCCACA 3170  
721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
3171 GAATCTCTTCTTCATCTGTTTAAACACACATATACCTGGAGCTACTTTTATTAACAAGAA 3230  
741 AspGlnGluValTySerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
3231 GACCAACAACTTGTPATAGTTCGCTTTCAAGGACATGGACAAATTTTCAGGTTTGTGT 3290  
761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyArgValSerMetThrThr 780  
3291 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCTATTTATGGTGTTCATGACGACT 3350  
781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
3351 TTGGAAGACGTATTTTAAAGCTAGAAATTTGAAGCAGAAATTTGACCAAGCAGATATATAG 3410  
801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
3411 GTATTTACTCAGCGCCATGGAGGAAGAAATGGATTCAAATCTTTTGTAGAAATGGA 3470  
821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
3471 CACAGCTTACTTATCTTCTGAAACCAAGGCTTCTAGTACGACACCATGAGCCCTTGG 3530  
841 LysGlnGlnMetTyThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
3531 AAACCAACAGATGATACATAGCAAGATTTCTATTTCTTACCTTGAACGTGAAGTAAA 3590  
861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
3591 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 3650  
881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTy 900  
3651 TTGGTTCATCACTCTTTTAAATATGCTGTGGTTTCCCATCAAACTTGTTCAGACTTATAT 3710

QY 901 PheLeuLysProGlyAspLysProHisLysTyLysThrSerLeuLeuLeuGlnAsnSer 920  
3711 TTTCTAAAACCTGGAGACAAACACATAAATACAAAACAAAGCTCTGCTTCTTCAAAATTCCT 3770  
QY 921 AlaAspSerAspIleSerAspLeuIleSerPheThrSerGlnAsnIleMetValThr 940  
3771 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGACATATAATGGTGACG 3830  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
3831 ATGATTAATAGACAGTGACTATCTATCCGTGGCTCCCATAGTCGGCTTTAAATGTGATG 3890  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
3891 CATTCAAAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTACTATGGTTATTCTTTA 3950  
QY 981 ProIleLeuValAsnIleleSerAsnTyLysLeuTyHisLeuAsnValThrGluThr 1000  
3951 CCTATATTAGTAATATCATTAAGTACTACTATCTTATCATTTAAATGTGCTGAACCC 4010  
QY 1001 IleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
4011 ATCCAGATCTGGAGTACCCTTCTTCAAGAAATTTACTGATATAGTTTTTAAATTTAG 4070  
QY 1021 LeuTyPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaValMetProProTyPheAla 1040  
4071 CTGTATTTTCAAGCAGCTTGTCTTGAATCAATTTGTACTGCAATGCCACTTACTTTGCC 4130  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyThrGlnLeuLysLeuSerGlyLeu 1060  
4131 ATGGAAATTCAGAGAAATCATAGATCAAGCTTATCTCAACTTAAACTTTCAGGCTTT 4190  
QY 1061 LeuProSerAlaTyTrpIleGlyGlnAlaValValAspIleProLeuPheIleIle 1080  
4191 TTGCCATCTGCATATTGGATTGACAAAGCTGTTGTGTATATCCCTTATTTTATCAT 4250  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyGlyLeuTyPheTyThrVal 1100  
4251 CTTATTTTGTGCTAGGAAGCTTACTGGCATTTTCAATTATGGATTAATTTTTTATCTGTA 4310  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyValProSerValIleLeuPheThr 1120  
4311 AAGTCTCTGCTGTGTTTGGCTTATGTTGTTTCCATCAGTATTCTGTTCAT 4370  
QY 1121 TyIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
4371 TATATTGCTTCTTTCACCTTTAAGAAAATTTTAAATACCAAGAAATTTTGGTCAATTA 4430  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTy 1160  
4431 TATCTGTGCAGCGTTGCTTGTATTGCAATCACTGAAATAACTTCTTTATGGGATAC 4490  
QY 1161 ThrIleAlaThrIleLeuHisTyAlaPheCysIleIleIleProIleTyProLeuLeu 1180  
4491 ACAATTCACACTATCTTCTTATGCTTTTGTATCATCATCTCCATCTATCCACTTCTA 4550  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
4551 GGTTCCTGATTTCTTTCAATAAGATTTCTTGAAGAAATGACGAAAAAATTTGGACACC 4610  
QY 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyTrpLeuGlnCysValLeu 1220  
4611 TATAATCCATGGGATAGGCTTTCAGTAGCTGTTTATATATCGCTTACCTGCTGCTGACTG 4670  
QY 1221 TrpIlePheLeuLeuGlnTyTyGluLysValTyGlyGlyArgSerIleArgLysAsp 1240  
4671 TGGATTTTCTCTTACAAATCTATGAGAAAAAATTTGGAGCAGATCAATAAGAAAAGAT 4730  
QY 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
4731 CCTTTTTCAGAAACCTTTCAACGAAGCTTAAAAATAGGAAGCTTCCAGAACCCACCCAGAC 4790



IKRKLCPALSMGLNPQITLLDEPSTGMDPKAQKQHWRAIRTAFTAFKNRKRAAILTHHNE  
 EAEAVCDVAINVSGQRICITGVQLXSKFGKYPLEIKLQKWIENLENDRLQREIQY  
 IFPNASRQESSTLAIKIPKEDYQSLSQSPFKLEAKHAFIIEEYFSQATLEQVY  
 BLTKEEBEDNSCGTLNLSLWERTQEDRVVF"

## ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 6525  
 Score: 8422.00 Matches: 1641  
 Percent Similarity: 99.94% Conservativity: 0  
 Best Local Similarity: 99.94% Mismatches: 1  
 Query Match: 99.95% Indels: 0  
 Gaps: 0

US-10-090-458-5 (1-1642) x AY028897 (1-6525)

QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20  
 Db 1011 ATGTCACCTGCAATTAGGAGGTAGGAGTTGGAGACACACAGAACACTTCTACTGAAG 1070  
 QY 21 AsnTyrLeuIleLysCysArgThrIleLysSerValGlnGluIleLeuPheProLeu 40  
 Db 1071 AATTACTTAAATTAATGACAGAACCAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 1130  
 QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
 Db 1131 TTTTITTTTATTTTGTGTTAATTAATTAGCATGATGCATCCAAATAAGAAATATGAAGA 1190  
 QY 61 ValProAsnIleGluLeuAsnProValAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
 Db 1191 GTGCGCTAATATAGAACTCAATCTATGACCAAGTTTACTCTTTCTAATCTAATTTCTTGA 1250  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 Db 1251 TATATCTCAGTGCACTAATATACAGACAGCATCATGCAAGAAAGTCTCTACTGATCATCTA 1310  
 QY 101 ProAspValIleIleThrGluIleThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 Db 1311 CCTGATGTCATATTAATGAGAAATATACAAATGAAGAAAGAAATGTTAAACATCCAGTCTC 1370  
 QY 121 SerLysProSerAsnPheValGlyValPheLysAspSerMetSerTyrGluLeuArg 140  
 Db 1371 TCTAAGCGAGCAACTTTGTAGTGTGTTTCAAGACTCCATGCTCTATGAATTCGT 1430  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 Db 1431 TTTTITTTCTGATGATGATCCAGTATCTCTCTATTTATGATGATTCAGAGCTGGCTGTTC 1490  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 Db 1491 AAATCATGTGAGCTGCTCAGTACTGCTCTCAGTTTCCAGTTTTCACAGTTTTCACAGCATCCATA 1550  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
 Db 1551 GATGCTGCCATTATACAGTTGAAGACCAATGTTTCTTTGAAGAGAGCTGGAGTCAACT 1610  
 QY 201 LysAlaValIleMetGlyGluThrAlaValAlaGluIleAspThrPheProArgGlyVal 220  
 Db 1611 AAAGCTGTATTATGGAGAAACTGCTGTGTAGAAATAGATACCTTTCCCGGAGGAGTA 1670  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 Db 1671 ATTTTAATATACCTAGTATAGCAATTTTCCACTTTTGGATACCTTTTGGCAATTCATATC 1730  
 QY 241 ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 Db 1731 GTAGCAGAAAAAAGAAAAATAAAGAAATTTTAAAGATTAATGGGACTTTCATGATACT 1790  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 Db 1791 GCCTTTTGGCTTTCTCTGCTTCTATATACAGTTTAATTTTCTTATGTCCTCTTCT 1850  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300

Db 1851 ATGGCAGTCATTCGCACAGCTTCTTTGTTATTTCTCAAGTAGCAGCATTCGTGATATT 1910  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 Db 1911 CTGCTTTTCTTCTTATGGATTATCATCTGTATTTTGTCTTAAATGCTGACACTCTT 1970  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 Db 1971 TTTAAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTGTGCTGCTTTTGATTT 2030  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 Db 2031 ATTGGCTTATCATTAATCTCATAGAAAGTTTTCCAAATCGTTAGTGTGGCTTTTCAGT 2090  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 Db 2091 CTTTCTGTCACGTACTTTTGTGATGTGTATGTCACAGGTGCATCATTTGAAGATTTT 2150  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 Db 2151 AATGAAGGTGCTTCAATTTTCAAAATTTGACTGCAGGCCCATATCTCTAATATTACAAT 2210  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 Db 2211 ATCATGCTCACACTTAATAGTATATCTATGCTCTTGGCTGTCTATCTTGTATCAAGTC 2270  
 QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 Db 2271 ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTATATATTTCTGAAGCTTCATATGG 2330  
 QY 441 SerLysSerLysArgAsnTyrGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 Db 2331 TCAAGAGTAAAGAAATATGAGGAGTTATCAGAGGGCAATGTTAATGGAATATTAGT 2390  
 QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
 Db 2391 TTTAGTGAATATTATGAGCCAGTCTTCTCAGAAATTTAGGAAGAAGACCATAGAATT 2450  
 QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
 Db 2451 AGTGGTATTTCAGAAACATACAGAAAGAGGGTGAATGTGGAGCTTTGAGAATTTG 2510  
 QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 Db 2511 TCAATTTGACATATATGAGGGTCAGATTACTGCTTACTTTGGCCACAGTGGACAGGAAG 2570  
 QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540  
 Db 2571 AGTACATTGATGAATATTCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 2630  
 QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 Db 2631 TATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGACAGAAATAATGATGGCATT 2690  
 QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 Db 2691 TGTCCACAGTTAGATATACATTTGATGTTTGCACAGTAGAAGAAATTTATCAATTTG 2750  
 QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 Db 2751 GCTTCAATCAAGGGATACAGCAACAATAATAACAAGAGTGCAGAGGGTTTACTA 2810  
 QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 Db 2811 GATTTAGACATGCAGACTATCAAGATACCAAGCTAAAAAATAAGTGGTGGTCAAAA 2870  
 QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuAspGlu 640  
 Db 2871 AGAAAGCTGTCTTAGGAATTTGCTTCTGGAAACCCAAAGATATCTGCTGCTAGTAGAA 2930  
 QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660

Db 2931 CCAACAGCTGGAATGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAATAACAGA 2990  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 2991 AAAGCCAAATCGGGTGACAGTGTTCAGTACTCAATTCATGGATGAAGCTGACATTTTGCA 3050  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 3051 GATAGAAAAGCTGTGATATACAAAGGAATGCTGAAATGTTGGTTCTTCAATGTTCTCTC 3110  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 3111 AAAAGTAAATGGGGATCGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACA 3170  
QY 721 GluSerLysSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db 3171 GAATCTCTTCTTCACTGTTTAAACAAACATATACCTGGAGCTACTTTATTACACAGAAAT 3230  
QY 741 AspGlnGluLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 3231 GACCAACAACCTGTGTATAGCTTGCTTTTCAAGGACATGGACAAATTTTCAGGTTTGTGT 3290  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 3291 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCAATTTCTATGGTGTTCATGACGACT 3350  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 3351 TTGGAGACGTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTCACCAAGCAGATATAGT 3410  
QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 3411 GTATTACTCAGCAGCCACTGGAGAGAAATGGATTCAAAATCTTTTGTATGAATGAA 3470  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 3471 CAGAGCTTACTTATTCTTTCTGAAACCAAGGCTTCTTAGTGAGCACCATCAGGCTTTGG 3530  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db 3531 AAACACAGATGTATACATAGAAAGTTTCATTTCTTACCTTGAACGTTGAAGTAA 3590  
QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 3591 TCAGTGAGATCAGTGTGCTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTT 3650  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 3651 TTGGTTTCATCTCTTTTAAATAATGCTGTGGTTCCCATCAAACTTGTTCAGACTTATAT 3710  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 3711 TTTCTAAAACCTGGAGACAAACCAATAAATACAAACAAGTCTGCTTCTTCAAAATCT 3770  
QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 3771 GCTGACTCAGATPATCAGTGATCTTATTAGCTTTTTCACAAGCCAGAACATATGCTGACG 3830  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 3831 ATGATTAATGACAGTGACTATGATATCCGTGGCTCCCAATAGTGGCTTTTAAATGTGATG 3890  
QY 961 HisSerGluLysAspTyrValPheAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 3891 CATTCAGAAAAGACTATGTTTTTTCAGCTGTTTTTCAACAGTACTATGTTTATTTCTTTA 3950  
QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 3951 CCTATATTAGTGAATATCATTAGTAACACTATCTTTTATCATTTTAAATGTGACTGAAACC 4010  
QY 1001 IleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 4011 ATCCAGATCTGGAGTACCCCACTCTTTTCAAGAAATTAAGTATAGTTTTTAAAAATTGAG 4070

QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 4071 CTGTATTTTTCAGCAGCTTTGCTTGGAAATCATTTGTACTGCAATGCCACCTTACTTTGCC 4130  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 4131 ATGGAAATCGAGAGAATCATAGATCAAAAGCTTATACTCAACTTAAACTTTCAGGTCTT 4190  
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleIle 1080  
Db 4191 TTGCCATCTGCATATTGGATTGGCAAGCTGTTGTGATATCCCTTATTTTATCATT 4250  
QY 1081 LeuIleLeuMetLeuGlySerLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 4251 CTATTTTGTATGCTAGGAGCTTACTGGCATTTTCATTATGGATATATTTTATACTGTA 4310  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 4311 AAGTCTCTGCTGCTGTTTTTTCCTTATTTGGTTATGTTCCATCAGTATTCTGTTCAC 4370  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 4371 TATATTGCTTCTTTTCCCTTTAAGAAAATTTTAAATACCAAGAAATTTTGGTCAATTATC 4430  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 4431 TATTCTGTGGCAGCTTGCTTGTATTGCAATCACTGAAATAACTTCTTTATGGGATAC 4490  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 4491 ACATTTGCAACTATTCTTCATTATGCTTTTGTATCATTCATTCATTCATTCATCTTA 4550  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
Db 4551 GGTGCTGCTGATTTCTTTTCAAAAGATTTCTTGGAGAATGTACGAAAAAATGTGGACACC 4610  
QY 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 4611 TATAATCCATGGATAGGCTTTTCAGTAGCTTTTATATCGCTTACTCTGCAAGTGTACTG 4670  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 4671 TGGATTTTCTCTTACATACTATGAGAAAAAATATGGAGGCAGATCAATTAAGAAAAGAT 4730  
QY 1241 ProPheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 4731 CCCTTTTTCAGAAAACCTTTCAACGAAGTCTAAAAATAGGAAGCTTCCAGAACCCAGAC 4790  
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 4791 AATGAGGATGAAGATCAAGATGTCAAAAGCTGAAAGACTAAAGGTCAAGAGCTGATGGT 4850  
QY 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 4851 TGCAGTGTGTGAGGAGAAACCATCCATTATGGTCAGCAATTTGTCATAAAGATATGAT 4910  
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysValLysValAlaThrLysTyrIleSer 1320  
Db 4911 GACAAAGAAAGATTTTCTTTTCAAGAAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 4970  
QY 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
Db 4971 TTCTGTGTGAAAAAGAGAGAGATCTTAGGACTATTGGGTCCAAATGGTCTCTGCGAAAAGC 5030  
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 5031 ACATTTATTATATTTCTGGTTGGTGATATTGAACCAACTTCAGGCAGAGTATTTTAGA 5090  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 5091 GATTATTTCTCAGAGCAAGTGAAGATGATGATTCATCTCACTGAAGTGTATGGGTACTGCTCT 5150

QY	1381	GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyxGlyAla	1400
Db	5151	CAGATAAACCCCTTGTGGCCAGATACATACATTGCAGGAACATTTTGAATTAATTTAGGAGCT	5210
QY	1401	ValIysGlyMetSerAlaSerAspMetIysGluValIleSerArgIleThrHisAlaLeu	1420
Db	5211	GTCAAGGAATGAGTCAAGTACATGAAAGAAAGTCATAAGTCGAATCAACATGCACCTT	5270
QY	1421	AspLeuIysGluHisLeuGlnIysThrValIysLysLeuProAlaGlyIleIysArgIys	1440
Db	5271	GATTTAAAGAACATCTTCAGAGACTGTAAAGAACTACCTGCAGGAATCAACGGAAG	5330
QY	1441	LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer	1460
Db	5331	TTGTGTTTTGCTCTAAGTATGCTAGGAACTCCATGATTTACTTTCTAGATGAACCATCT	5390
QY	1461	ThrGlyMetAspProIysAlaIysGlnHisMetTrpArgAlaIleArgThrAlaPheIys	1480
Db	5391	ACAGGTATGATCCCAACCAACCAACGACATGTGGCGAGCAATTCGATGCTATTAA	5450
QY	1481	AsnArgLysArgAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCys	1500
Db	5451	AACAGAAAGCGGCTGCTATTCTGACCACTCACTATATGGAGGAGCGAGGCTGCTCTGT	5510
QY	1501	AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu	1520
Db	5511	GATCGAGTACCTATCATGTGCTGCGGAGTTAGATGTATCCGAAACAGTACACATCTA	5570
QY	1521	LysSerLysPheGlyGlyGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn	1540
Db	5571	AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTAATTAAGAGCTGAGTAAAC	5630
QY	1541	LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln	1560
Db	5631	CTAGAAGTACAGCGCTTCAAGAGAAATTCAGTATATTTCCCAATGCAAGCGTCAG	5690
QY	1561	GluSerPheSerSerIleLeuAlaTyrIleProLysGluAspValGlnSerLeuSer	1580
Db	5691	GAAAGTTTTCTTCTATTGCTTATATAAATTCCTAAGGAAGATGTTCAAGTCCCTTCA	5750
QY	1581	GlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe	1600
Db	5751	CAATCTTTTTTAAGCTGGAAGCAATCAACATGCTTTTGCATTAAGATATAGCTTT	5810
QY	1601	SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp	1620
Db	5811	TCTCAGCAACATTCGAACAGCTTTTGTAGAACTCACTAAGAACAGAGAGGAAGAT	5870
QY	1621	AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal	1640
Db	5871	AATAGTTGTGGAACCTTTAAACAGCACCTTTGCTGGGAACGAAACAAAGAGATAGATA	5930
QY	1641	ValPhe	1642
Db	5931	GTAATT	5936
RESULT 5			
LOCUS	HSA275973		
DEFINITION	Homo sapiens mRNA for ATP-binding cassette protein of the (ABCA subfamily).		
ACCESSION	AJ275973		
VERSION	AJ275973.3	GI:22080663	
KEYWORDS	ABCA subfamily; ATP-binding cassette protein.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1		
AUTHORS	Petry, F., Kotthaus, A. and Hirsch-Ernst, K.I.		
TITLE	Molecular cloning and tissue distribution of a novel ATP-binding cassette (ABC) transporter belonging to the subfamily ABCA		
JOURNAL	Unpublished		

REFERENCE	2		
AUTHORS	Hirsch-Ernst, K.I.		
TITLE	Direct Submission		
JOURNAL	Submitted (02-MAR-2000) Hirsch-Ernst K.I., Department of Toxicology, University of Goettingen, Robert-Koch-Strasse 40, Lowery Saxony, D-37075, GERMANY		
REMARK	revised by [3]		
REFERENCE	3 (bases 1 to 5096)		
AUTHORS	Hirsch-Ernst, K.I.		
TITLE	Direct Submission		
JOURNAL	Submitted (20-NOV-2001) Hirsch-Ernst K.I., Department of Toxicology, University of Goettingen, Robert-Koch-Strasse 40, Lowery Saxony, D-37075, GERMANY		
COMMENT	On Aug 1, 2002 this sequence version replaced gi:17046099.		
FEATURES	Location/Qualifiers		
source	1..5096		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	66..4994		
	/tissue_type="testis"		
	/function="ABC transporter"		
	/notes="belongs to the ABCA subfamily"		
	/codon_start=1		
	/product="ATP-binding cassette protein"		
	/protein_id="CA393535.3"		
	/db_xref="GI:22080664"		
	/db_xref="GOA:Q9NYL4"		
	/translation="MSTAIRREVWVRQRTLLKNYLKCRKTKSSVQBIPLPFLF WILISMPHPKPKYKVEPNIELNPMDFKTLNLIIGLYPTVNTIMOKVSTDLHPD VITEBYTKEMKLTSSLSKSPNFVGFVDFKMSYELRFPDMIPVSSITMDSRAGCS KSCERAQYWSGGFTVLOASIDAALICLNTNLSLKESTKAVIMGETAVVIDTPPR GVILYLVAFSPFGYFLAHIVAEKEKIKFELKIMGLADTAFWLSWLLYTSILFL MSLLMAVITASLIPQSSSIVIFLLFLYGLSSVFFALMLTLPFKSKHGVIFEV TVAFGIGMILILIESPFSKLVLPSPFCHCTVIGIAQVHLEDFNEGASFNLTAG PPLIITLMLTNSFYVLLAYLDQVIPGFEGLRSLYFLKFSYMSKSNYBEL SEGNVNGTSFSEIIPVSSEFVGEARISGTYKTKYKKNALRNLSFDIYEGQ IYALLGHSCTGKSTLMNLCGLCPSPDGPAISVYHRVSDIEMFAKMGICPQDII HDVLVVENLSILASIKGPANNIIQVQKVLDDIMOTIKDNOAKLISGGQKRLS LGIAVLGNKILLIDPTAGMDPCSRHIVNLLKYRKANRVTVFSTHFWDEADILADR KAVISGMUKCVGSNFKSKWGIKGRSLNIDYIKYKATCSLSLKVQKHPGATLLQON DQOLVYQPKMDKPSGLFSAIDSHSNISVSYGVSMTLSDVPPNLSSTSKNKKIPEPPDNE YSVFTQPLEEEDSKSFDMEQSLISETKASLVSTMSLWKQNYTIAKFFFLK RSKSVRSVLLDILLIIPFTQIFMFLVHSFKNAVPIKLVDPDLYFLKPGDKPKHYAAT LILQNSADSDISLISFTSQNMIVMTINDSYVAVSHAALNVMGSEKDYFAAFV NTMAYSLPILAVNISYVLYHLNVTETIQIWTSPFQBITDVFKEIYFOAALIGL IYTMADPPYFAMENANENKIKAYTKLSGLSPSAWYGOAVDIPLFILILMLGSL LAFHYGLYTYVKFLAVVCLIGYVPSVILFYIASPTFKLINTKEFMSFYVSAAL ACIAITEITTFMGYTIATLHAFCIILIEPLGCLISFIKISKNKKNVDYTNPW DRLSVAVISPYLQCVLWIFLQYKKGYSIRKDPFNNLSSTSKNKKIPEPPDNE DEDEDVKAERLKVLMGQCCCEKPSMVSNLHKEYDDKDFLLSRKVKVATKYIS FCVKGEILGLLPGNAGKSTIINILVGDIEPTSGQIFLGDYSSSEDDSLKCMGY CPQINLPDPTLQEHFEYTGAVKMGASDMKEVISRIITHALDLKHLQKTKVCLPAG IKRKLCLFALSMGLNPQITLLDPSGMDPKAKHWMRAIRTAFAKKNRAAILUTHTYME EBAVCDRVALWVSGQLRCITQVHLKSKGKGYELEKLKDWIENLEVDRLQREIQY IPFNASRQSFSSILAIKIPKEDVQSLSQSFPLEAKHAPALEIYFSFQATLEQVFV ELTKQEEDNSCGTLNSTLWWTQEDRVF"		
Alignment Scores:			
Pred. No.:	0	Length:	5096
Score:	8421.00	Matches:	1640
Percent Similarity:	100.00%	Conservative:	2
Best Local Similarity:	99.88%	Mismatches:	0
Query Match:	99.94%	Indels:	0
DB:	9	Gaps:	0
ORIGIN			
	US-10-090-458-5 (1-1642) x HSA275973 (1-5096)		
QY	1	MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrIleLeuLeuLys	20
Db	66	ATGTCACCTGCAATTAGGAGGTAGGAGTTTGAGACACAGACACACTTCTACTGAAG	125

QY 21 AsnTyrLeuIleLysCysArgThrIlyLysSerValGlnGluIleLeuPheProLeu 40  
DB 126 AATTACTTAATTAATGCAGACCAACAAAGAGTAGTGTTCAGAAATTTCTTTTCCACTA 185  
QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
DB 186 TTTTITTTTATTGGTTAATTAATTAATAGCATGATGCATCCAAATTAAGAAATATAGAGAA 245  
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
DB 246 GTGCCCTAATATAGAACTCAATCTATGGACAAGTTTACTCTTCTTAATCTAATTTCTTGGAA 305  
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
DB 306 TATACTCCAGTGACTAATATTACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA 365  
QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
DB 366 CCTGATGTCATTAATTAATGAAATATACAATGAAAGAAATGTTAACTCCAGTCTC 425  
QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
DB 426 TCTAAGCCGAGCAACTTTGTAGTGTGTTTCAAGACTCCATGCTCTATGAATTCGT 485  
QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
DB 486 TTTTITTCCTGATGATGATTCAGTATCTTCTATTATATGGATTCAGAGCTGGCTGTTC 545  
QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
DB 546 AATCATGTGAGCTGCTCAGTACTGTGCTCCTCAGGTTTCAAGTTTACAGCATCCATA 605  
QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
DB 606 GATGCTGCCAATTATACAGTTGAAGCAATGTTCTCTTTGGAAGGAGCTGGAGTCACT 665  
QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
DB 666 AAGCTGTATTATGGGAAACTGCTGTGTGAGAAATAGATACCTTTCCCGAGGAGTA 725  
QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
DB 726 ATTTTAAATACCTAGTTATAGCATTTTCACCTTTTGGATATCTTTTGGCAATTCATATC 785  
QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
DB 786 GTAGCAGAAATAAGAAATAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 845  
QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
DB 846 GCCTTTTGGCTTTCCTGGGTTCTTCTATATACAAGTTTAAATTTTCTTATGCTCCTTCTT 905  
QY 281 MetAlaValIleAlaIleAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
DB 906 ATGGCAGTCATTCGACAGCTCTTTTGTATTATTCCTCAAGTAGCAGCATTTGTGATATT 965  
QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
DB 966 CTGCTTTTTCCTTTAAGATTATCATCTGTAATTTTGTGTTTAACTGCTGACACCTTCT 1025  
QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
DB 1026 TTTTAAATAATCAAAACATGTGGGAATAGTTGAATTTTGTGTTTGTGCTTTTGGATT 1085  
QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
DB 1086 ATTGCCCTTATGATTAATCTCTAAGAAAGTTTCCCAATCTGTTAGTGTGCTTTTCAGT 1145  
QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
DB 1146 CCTTTCTGTCACGTACTTTTGTGATTTGATTTGACAGGTCATGCATTTAGAGAAATTT 1205

QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
DB 1206 AATGAAGTGTCTTCATTTTCARATTTGACTGCAGGCCCATATCTCTTAATTATACAAAT 1265  
QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
DB 1266 ATCATGCTCACACTTAATAGTATATTCTATGCTCTCTGGCTGTCTATCTTGATCAAGTC 1325  
QY 421 IleProGlyGluPheGlyLeuArgAspSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
DB 1326 ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATATTTCTGAAGCTTCATATTGG 1385  
QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
DB 1386 TCAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGCAATGTTAATGGAAATATTAGT 1445  
QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
DB 1446 TTTAGTGAATATTATGAGCCAGTTTCTTCAGAAATTTGTAGAAAGAACCCATTAAGAT 1505  
QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
DB 1506 AGTGTATTTCAGAAAGACATACAGAAAGAGGTGAAATGTGGAGCTTTGAGAAATTTG 1565  
QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
DB 1566 TCATTTGACATATATGAGGTCAGATTACTGCTTACTTGGCCACAGTGGAAACGGAAG 1625  
QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
DB 1626 AGTACATTGATGAATATTCTTTGTGACTCTGCCACCTTCTGATGGTTTGCATCTATA 1685  
QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
DB 1686 TATGSCACACAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATGGCAAT 1745  
QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
DB 1746 TGTCCACAGTTAGATATACACTTTGATGTTTTCAGACAGTAGAAGAAATTTATCAATTTG 1805  
QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
DB 1806 GCTTCAATCAAGGGATACAGCCAAATATAATACAAGAGTGCGAAGGTTTTCATA 1865  
QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
DB 1866 GATTTAGACATGACAGACTATCAAGATAACCAAGCTAAAAAATTAAGTGTGTCAAAAA 1925  
QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
DB 1926 AGAAAGCTGTCTATTAGGAATTTGCTTCTTGGCAACCCAAAGATATCTGCTAGATGAA 1985  
QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
DB 1986 CCAACAGCTGGAAATGGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAATACAGA 2045  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
DB 2046 AAAGCCAATCCGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTTGA 2105  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
DB 2106 GATAGAAAGCTGTGATATACAGGAATGCTGAAATGTGTGGTTCTTCAATGTTCCCTC 2165  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 2166 AAAAGTAAATGGGGATCGCTACCCCTGAGCATGTACATAGACAAATATTGTGCGACA 2225  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
DB 2226 GATCTCTTTCTTCTCTGTTAAACACACATATACCTGGAGCTACTTTATTACACAGAAAT 2285  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760



Db 2286 GACCAACAACCTTGCTGTAAGCTTGCCTTCAAGGACATGGCAAAATTTTCAGGTTTGTT 2345  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerThrGlyValSerMetThr 780  
Db 2346 TCTGCCCTAGACAGCATTCAAATTTGGGTGTCATTTCTTATGGTGTTTCATGACGACT 2405  
QY 781 LeuGluAspValPheLeuLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2406 TTGGAAGACGCTATTTTAAAGCTAGAAGCTGAAGCAGAAATTTGACCAAGCAGATTAAGT 2465  
QY 801 ValPheThrGlnGlnProLeuGluGluIleuWetAspSerLysSerPheAspGluMetGlu 820  
Db 2466 GTATTATTACTGACGACCACTGGAGAGAAATGGAATCAAAATCTTTTGATGAATGAA 2525  
QY 821 GlnSerLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
Db 2526 CAGAGCTTACTTATCTTCTTGAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGG 2585  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgLysSerLys 860  
Db 2586 AAACAACAGATGTAACAAATAGCAAGATTTCAATTTCTTACCTTGAACCGTGAAGTAA 2645  
QY 861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2646 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTT 2705  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 2706 TTGGTTCATCACTCTTTTAAAAATGCTGTGTGTTCCCATCAAACTTGTTCAGACTTAT 2765  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 2766 TTTCTAAACCTGGAGACAACCAATTAATACAAACAACTCTGCTTCTTCAAAATCT 2825  
QY 921 AlaAspSerAspIleSerAspLeuLeuSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 2826 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGAACATTAATGGTGAG 2885  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 2886 ATGATTAAATGACAGTGACTATGATCCGGTGGCTCCCATAGTGGCGCTTTAAATGTGATG 2945  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 2946 CATTCAGAAAGACATGCTTTTTCAGCTGTTTTCACAGTACTATGTTTATCTTTA 3005  
QY 981 ProIleLeuValAsnIleLeuSerAsnTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 3006 CCTATATTAGTGAATATCATTAAGTACTATCTTTATCATTTAAATGTGACTGAAC 3065  
QY 1001 IleGlnIleTyrSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 3066 ATCCAGATCTGGAGTACCCATCTTTCAAGAAATTAATGATATAGTTTAAATGAG 3125  
QY 1021 LeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 3126 CTGTATTTTCAAGCAGCTTGTGGAATCAATGTTTACTGCAATGCCACCTTACTTTGCC 3185  
QY 1041 MetGluAsnAlaGluAsnHisLysValIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3186 ATGGAATTCGAGAGAAATCATAGATCAAGCTTATTAACAACTTAACCTTCACTTCAGGCTT 3245  
QY 1061 LeuProSerAlaTyrTrpIleGlnAlaValValAspIleProLeuPhePheIle 1080  
Db 3246 TTGCCATCTGCATTTATTTGGATTGGACAAGCTGTGTTGATATCCCTTATTTTATCAT 3305  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3306 CTTATTATTGATGCTAGGAAGCTTATTGGCATTTTCAATTATGGATTAATTTTATTAAGTGA 3365  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120

Db 3366 AAGTTCTCTGCTGCTGGTTTTCCTTATGTTGTTAAGTTCATCAGTATTTCGTGCTACT 3425  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTyrSerPheIle 1140  
Db 3426 TATATTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3485  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3486 TATCTGTGGCAGCTTGGCTTGTATTCATCACTGAAATAAATCTTCTTATGGGATAC 3545  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 3546 ACATTTGCACTATTCTTTCATTTATGCCCTTTGTATCATCAATTCATCTCACTTCTA 3605  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerThrLysAsnValArgLysAsnValAspThr 1200  
Db 3606 GGTGCTGCTGATTTCTTTCATAAAGATTTCTTGAAGAATGTACGAAATAATGTGACACC 3665  
QY 1201 TyrAsnProTyrAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 3666 TATATCCATGGATAGGCTTTCAGTAGCTGTATATCGCCTTACTCGCAGTGTGACTG 3725  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 3726 TGGATTTTCTCTTCAATATCTATGAGAAAAATATGGAGGCAGATCAATAAGAAAGAT 3785  
QY 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 3786 CCTTTTTTCAGAAACCTTTCAACGAAAGTCTAAATATGGAAGCTTCCAGAACCCAGAC 3845  
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 3846 AATGAGATGAAGATCAGATGCTCAAGCTGAAAGACTAAAGCTCAAGAGCTGATGGT 3905  
QY 1281 CysGlnCysGlyGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 3906 TGCACGTGTGTGAGAGAAACCATTCATTTGGTCAGCAATTTGCATTAAGAATATGAT 3965  
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer 1320  
Db 3966 GACAAAGAAATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 4025  
QY 1321 PheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
Db 4026 TTCTGTGTGAAAGAGAGAGATCTTAGACATTTGGGTGCAATGGTCTGCGAAGAC 4085  
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 4086 ACATTTATTAATTTCTGTTGGTGATTTGAACCAACTTCAGGCCAGATATTTTAGGA 4145  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 4146 GATTATTTCTCAGAGTCAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTACTGCTCT 4205  
QY 1381 GlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
Db 4206 CAGATAAACCTTTTGGCAGAGATCACTATTCAGCAAGACATTTTGAATTTATGAGCT 4265  
QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db 4266 GTCAAAGGAATGAGTCAAGTGCATGAAGAAAGTCAATAGTGAATAACACATGACCTT 4325  
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
Db 4326 GATTTAAAGAACATCTTCAGAGATCTTAAGAAACTACCTGACGAATCAACCAAG 4385  
QY 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
Db 4386 TTGTTTCTCTCTAAGTATGCTAGGAAATCTCTCAGATTAATTTTGTAGTAGAACCATCT 4445  
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
Db 4446 ACAGGTATGATCCCAAGCCAAACAGCACATGTGGCGAATTCGGACTGATTTAA 4505

```

QY 1481 AsnArgLysArgAlaAlaIleuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500
Db 4506 AACGAAAGCGGGTCTGCTATTCTGACCACTCACTATATGGAGGAGGAGGCTGTCTGT 4565

QY 1501 AspArgValAlaAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520
Db 4566 GATCGAGTAGCTATCATGCTGTCTGGGCGAGTTAGATGTTATCGAACAGATACACATCTA 4625

QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTyrIleGluAsn 1540
Db 4626 AAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTTAAATTTGAAGGACTGCGATAGAAAAC 4685

QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560
Db 4686 CTAGAAGTAGACCGCTTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCCGTCAG 4745

QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580
Db 4746 GAAAGTTTTCTCTATTTTGGCTTATAAAATTTCTTAAGGAAGATGTTTCAGTCCCTTTCA 4805

QY 1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600
Db 4806 CAATCTTTTTTAAAGCTGGAAGAACTAAACATGCTTTTGGCATTGAAGAATATAGCTTT 4865

QY 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620
Db 4866 TCTCAGCAACATTTGGAACAGCTTTTGTAGAACTCTAATAAGAACAGAGGGAAGAT 4925

QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyrTrpGluArgThrGlnGluAspArgVal 1640
Db 4926 AATAGTTGTGGAACCTTTAAACAGCACACTTTGTGGGAAACGAAACACAAAGAGATAGAGTA 4985

QY 1641 ValPhe 1642
Db 4986 GTATTT 4991

RESULT 6
AX417824
LOCUS AX417824 4929 bp DNA linear PAT 18-JUN-2002
DEFINITION Sequence 1 from Patent WO0231147.
ACCESSION AX417824
VERSION AX417824.1 GI:21522942
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Hu.Y. and Nepomnichy,B.
TITLE Human transporter proteins and polynucleotides encoding the same
JOURNAL Patent: WO 0231147-A 1 18-APR-2002;
LEXICON GENETICS INC (US)
FEATURES
source
1. 4929
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4929
Score: 8395.00 Matches: 1637
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 5
Query Match: 99.63% Indels: 0
DB: 6 Gaps: 0

US-10-090-458-5 (1-1642) x AX417824 (1-4929)

QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20
Db 1 ATGTCCACTGCAATTAGGAGGATAGGATTTGGAGACAGACCAAGCACTTCTACTGAAG 60

```

```

QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu 40
Db 61 AATTACTTAATTAATGAGAACCAAAAGAGTAGTGTTCAGAAATCTTTTCCACTA 120

QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 121 TTTTITTTTATTTTGGTTATATTAATTTAGCATGATGATCCAAATTAAGAAATATGAGAA 180

QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleGly 80
Db 181 GTGCTTAATATAGAACTCAATCCTATGGACAAGTTTACTCTTTCTAATCTAATCTTGA 240

QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
Db 241 TATATCTCCAGTCACATAATTTACAAGCAGCATCATGCGAAAGTGTCTACTGATCATCTA 300

QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
Db 301 CCTGATGTCATAATTTACTGAGAATATACAATGAAAGAAATCTTAACATCCAGTCTC 360

QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
Db 361 TCTAAGCCGAGCAACTTTGTAGTGTGTGTTTCAAAGACTCCATGCTCTATGAATCTGT 420

QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
Db 421 TTTTITTTCTGATATATGATCCAGTATCTTCTATATATATGATTCAGAGCTGGCTGTTC 480

QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180
Db 481 AAATCATGTGAGGCTGCTCAGTACTGTGCTCCTCAGGTTTACAGATTTTACAAGATCCATA 540

QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200
Db 541 GATGCTGCCATTTATACAGTTGAAGACCAATGTTTCTTTTGAAGAGCTGGAGTCAACT 600

QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
Db 601 AAAGCTGTATTATATGGGAGAACTGCTGTGTAGAAATAGATACCTTTTCCCGAGAGTA 660

QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
Db 661 ATTTTAAATATACCTAGTTATAGCATTTTACCTTTTGGATACTTTTGGCAATTCATATC 720

QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
Db 721 GTAGCAGAAAAGAAAAGAAAATTAAGAAATTTTAAAGATAATGGGACTTCATGATACT 780

QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280
Db 781 GCCTTTTGGCTTCCCTGGGTCTTCTATATAWACAAGTTTAAATTTTCTTATGTCCTTCT 840

QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300
Db 841 ATGGCAGTCATTCGACAGCTCTTTTGTATTCTCTCAAGTAGCAGCATTTGTGATATT 900

QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320
Db 901 CTGCTTTTTTCTCTTATGGAATATCATCTGTATTATTTTGTCTTTTAACTGTCACCTCT 960

QY 321 PheLysLysSerLysHisValGlyIleValGluPheValThrValAlaPheGlyPhe 340
Db 961 TTTTAAATATCAAAACATGTGGGAATAGTTGAAATTTTGTGTACTGTGGCTTTTGAATT 1020

QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360
Db 1021 ATTTGGCTTATGATTAATCTCATAGAAAATTTTCCCAATCGTTAGTGTGGCTTTTCAGT 1080

QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380
Db 1081 CCTTTCTGTCTACTGTTTGTGATTTGTGATTTGACAGGTCATGCATTTTGAAGATTT 1140

```

QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 DB 1141 AATGAAGGTGCTTCATTTTCAAAATTGACTCGAGGCCCATATCTCTAAATATTACAAAT 1200  
 QY 401 IleValLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 DB 1201 ATCATGCTCACACTTAATAGATATATCTATGCTCTCTGCTGCTATCTTGATCAAGTC 1260  
 QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 DB 1261 ATTCGAGGGAATTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTGG 1320  
 QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 DB 1321 TCAAAGACCAAAAGAAATATGAGAGATTATCAGAGGCAATGTTATATGGAATATTAGT 1380  
 QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
 DB 1381 TTTAGTGAATTTATTGAGCCAGTTCTTCAGAAATTTGTAGGAAAGAGCCATAAGAAAT 1440  
 QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
 DB 1441 AGTGTATTTCAGAACACATACAGAAAGAGGTGAAAATGTGGAGGCTTTGAGAAATTTG 1500  
 QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 DB 1501 TCATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCCACAGTGGACAGGAAAG 1560  
 QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540  
 DB 1561 AGTACATTGATGAATATTCTTTGTGGACTCTGCCACCTCTCTGATGGGTTTGGATCTATA 1620  
 QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 DB 1621 TATGACACAGAGTCTCAGAAATAGATGAATGTTTGAGCAAGAAATGATGGCAT 1680  
 QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 DB 1681 TGTCACAGTTAGATATACATCTTTGATGTTTGTGACAGTAGAAGAAATTTATCAATTTTG 1740  
 QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 DB 1741 GTTTCATCAAGAGGATACAGCCACCAATATATATACAAAGATGCGAAGGTTTACTA 1800  
 QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 DB 1801 GATTTAGACATCGACTATCAAGATAACCAAGCTAAATAAATTAAGTGTGGTCAAAA 1860  
 QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
 DB 1861 AGAAAGCTGTCAATTAGGAATTGCTGTTCTTGGGAACCCAAAGATACTGCTGATAGTAA 1920  
 QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTyrPheAsnLeuLysTyrArg 660  
 DB 1921 CCACAGCTGGAATGGACCCCTGTTCTCGACATATTGATGGAATCTTTTAAATACAGA 1980  
 QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 DB 1981 AAAGCCAAATCGGGTGACGTGTCAGTACTCATTTTCATGGATGAAGCTGACATCTTGCA 2040  
 QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 DB 2041 GATAGGAAGCTGTGATATCACAGGAATGCTGAAATGTGTTGTTGTTCTTCATGTTCTCT 2100  
 QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 DB 2101 AAAAGTAAATGGGGATCGCTACCGCTGACATGTATACAGCAAAATATGTTGCCACA 2160  
 QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
 DB 2161 GAAATCTTTCTTCACTGGTTAAACACATATATCTGGAGCTACTTTTATACACAGAAAT 2220  
 QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760

DB 2221 GACCAACAACTGTGTATAGCTTGCTTTCAAGGACATGGACAAATTTTCAGTTTGT 2280  
 QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
 DB 2281 TCTGCCCTAGACAGTCATTCAAATTTGGTGCGATTTCTTATGGGTTTCCATGACGACT 2340  
 QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
 DB 2341 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATTATAGT 2400  
 QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
 DB 2401 GTATTTACTCAGCAGCCACTGGAGGAGAAATGGAATCAAATCTTTGTGAAATGAA 2460  
 QY 821 GlnSerLeuLeuIleLeuSerGlnThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
 DB 2461 CAGAGCTTACTTATTTCTTTCTGAAACCAAGGCTCTCTAGTAGACCAACCATGAGCTTTGG 2520  
 QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
 DB 2521 AAACACAGATGATACATACAAAGTTTCATTTCTTACCTTGAAACGTGAAAGTAA 2580  
 QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
 DB 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTTTCACAGTTTCAGATTTTATGTT 2640  
 QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
 DB 2641 TTGGTTCACTCATCTCTTTTAAAATGCTGTGGTTCCTCATCAAACTTTGCCAGCTTATAT 2700  
 QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
 DB 2701 TTTCTAAAACCTGGAGACAAACACATAAATCAAAACAAGTCTGCTCTTCTCAAAATCT 2760  
 QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
 DB 2761 GCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAGCCAGACATATGTTGAGC 2820  
 QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
 DB 2821 ATGATTAATGACAGTGACTATGTATCCGTGGCTGCCCATAGTCGGCTTTAAATGTGTG 2880  
 QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
 DB 2881 CATTCAGAAAGAGACTATGTTTTCAGCTGTTTTCACAGPACTATAGTTTATCTTTA 2940  
 QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
 DB 2941 CCTATATTAGTGAATATCATTAAGTAACTATCTTTATCAATTTAAATGTGACGAAAC 3000  
 QY 1001 IleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
 DB 3001 ATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTTACTGATATAGTTTAAAAATTGAG 3060  
 QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAla 1040  
 DB 3061 CTGTATTTTCAGCAGCTTTGCTTGGAAATCATTGTTACTGCAATGCCACCTTACTTTGCC 3120  
 QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
 DB 3121 ATGGAAAAATCGAGAGAAATCATAGATCAAAAGCTTAACTCAACTTAACTTTCAGCTCT 3180  
 QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
 DB 3181 TTGCCATCTGCATATTGGATTGGACAGCTGTTGTTGATATCCCTTATTTTATCAAT 3240  
 QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheThrVal 1100  
 DB 3241 CTTATTTTGAATGCTAGGAAGCTTATTTGGCATTTCTATGGATATATTTTATATCTGTA 3300  
 QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120



QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu 40  
Db 315 AATTAACCTTAATTAATGAGAAACCAACCAAGAGTAGTGTTCAGGAATATCTTTTCCACTA 374  
QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysValGlnGlu 60  
Db 375 TTTTATTTATTTTGGTTAATATTAATATAGCATGATGCAATCCAAATAGAAATATGAAGA 434  
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
Db 435 GTGCGCTAATATAGAACTCAATCCTATGACCAAGTTTACTCTTCTTAATCTAATCTTTGGA 494  
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
Db 495 TATATCTCCAGTCACTAATATTAACAGCAGCATCATGCAAGAGTGTCTACTGATCACTA 554  
QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
Db 555 CCTGATGTCATAATTACTGAAGATATACAAATGAAAAAGAAATGTTAAATCCTAATCTTTGGA 614  
QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
Db 615 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAGAGCTCCATGCTCTATGAATCTTGT 674  
QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
Db 675 TTTTTCCTGATGATGATCCAGTATCTTCTATTTATATGGATTCAAGAGCTGGCTGTCA 734  
QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
Db 735 AAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTACAGTATTTTCAAGCATCCATA 794  
QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
Db 795 GATGCTGCCATTATACAGTTGAAGCAATGTTTCTCTTGGAGAGAGCTGGAGTCAACT 854  
QY 201 LysAlaValIleMetGlyGluThrAlaValAlaGluIleAspThrPheProArgGlyVal 220  
Db 855 AAAGCTGTTATTATGGAGAAACTGCTGTGTGAGAAATAGATACACTTTCCCGAGGAGTA 914  
QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
Db 915 ATTTTAATATACCTAGTTATAGCATTTTCCACTTTTGGATATCTTTTGGCAATTCATATC 974  
QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
Db 975 GTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 1034  
QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
Db 1035 GCCTTTTGGCTTTCCCTGGTTCTTCTATAWACAAGTTTAAATTTTCTTATGFCCTTCTT 1094  
QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300  
Db 1095 ATGGCAGTCATTCGACAGCTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTGATATT 1154  
QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
Db 1155 CTGCTTTTTCCTTTATGGAATATCATCTGTATTTTGTCTTTAATGCTGACACTCTT 1214  
QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
Db 1215 TTTTAAAAATCAAAACATGTGGGATAGTTGAAATTTTGTACTGTGGCTTTTGATTT 1274  
QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
Db 1275 ATTTGGCTTATGATAATCCTCATAGAAAGTTTTCCAAATCGTTAGTGTGGCTTTTCAGT 1334  
QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
Db 1335 CCTTTCTGTCAGTACTTTTGTGATTTGTTATGTCAGGTATGCAATTTAGAAATTTT 1394

QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
Db 1395 AATGAAGGTGCTTCATTTTCAAATTTGACTGAGGCCCATATCCTCTAATATTATTAACAAT 1454  
QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
Db 1455 ATCATGCTCACACTTAATAGTATATCTATGTCCTTCTGCTGCTCTATCTTGATCAAGTC 1514  
QY 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
Db 1515 ATTCCAGGGGAATTTGGCTTACGAGATCATCTTTATATATTTTCTGAAGCCTTCATATTGG 1574  
QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
Db 1575 TCAAGAGCANAAGAAATTTATGAGGATTTATGAGGGCAATGTTAATGGAATATTACT 1634  
QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
Db 1635 TTTAGTGAATATTATTGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAAGAACCCATAAGAAAT 1694  
QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyValAsnValGluAlaLeuArgAsnLeu 500  
Db 1695 AGTGGTATTCAGAGACATACAGAAAGAGGTGAAATGTGGAGCTTTGAGAAATTTG 1754  
QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
Db 1755 TCATTTGACATATATGAGGTCAGATTACTGCTTACTTTGGCCACAGTGGAAACAGGAAAG 1814  
QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
Db 1815 AGTACATTTGATGAATATTTCTTTGTGACCTTCTGCCACCTTCTGATGGGTTTGCACTATA 1874  
QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db 1875 TATGGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTGGCAAT 1934  
QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
Db 1935 TGTCACAGTTAGATATACACTTTTGATGTTTGTACAGTAGAAGAAATTTATCAATTTTG 1994  
QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
Db 1995 GCTTCAATCAAGGGATACAGCCCACTAATAATAACAAGAGTCAGAAAGGTTTACTA 2054  
QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys 620  
Db 2055 GATTTAGACATGACAGTATCAAAAGATAACCAAGCTAAAAAATAAGTGGTGTCAAAA 2114  
QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuAspGlu 640  
Db 2115 AGAAGCTGTCTAGGAATTTGCTTCTTGGAAACCAAGATACTGCTGTAGATGAA 2174  
QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyrArg 660  
Db 2175 CCAACAGCTGGAATGGACCCCTGTTCTCGACATATTGTATGGAAATCTTTTAAATACAGA 2234  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 2235 AAAGCCAAATCGGTGACAGTTCAGTACTCATTTTATGGAAGAGTGCATCTTGCA 2294  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2295 GATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTGTGGTCTTCTCAATGTTCCTC 2354  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2355 AAAAGTAAATGGGGATTCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGGCACA 2414  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
Db 2415 GAATCTCTTCTTCACTGGTTAAACAACATATACCTGGAGCTACTTATTATACAACAGAA 2474  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760

Db 2475 GACCAACAACCTTGTAAGCTTGGCTTTCAGGACATGGACAAATTTTCAGGTTTGTTT 2534  
Qy 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2535 TCTGCCCTAGACAGTCATTCAAATTTGGGTGGCAATTTCTTATGGGGTTTCCATGACGACT 2594  
Qy 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2595 TTGGAGAGCTATTTTAAAGCTAGAGTTGAGCGAGAAATTTGACCAAGCAGATTATAGT 2654  
Qy 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2655 GTATTACTCAGCAGCCACTGAGGAGAAAGATGGATTCAAAATCTTTTGTAGAAATGGAA 2714  
Qy 821 GlnSerLeuLeuLeuSerGlnThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
Db 2715 CAGAGCTTACTTATTCTTCTGAAACCAAGGCTCTCTAGTGAGCACCATGAGCCTTTGG 2774  
Qy 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db 2775 AAACAACAGATGATACAAATGCAAAAGTTTCATTTCTTTTACCTTGAAACGGTGAAGTAAA 2834  
Qy 861 SerValArgSerValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 880  
Db 2835 TCAGTGAGATCAGTGTGCTTCTGCTCTTTAAATTTTTCATGAGTTTCAGATTTTATGTTT 2894  
Qy 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 2895 TTGGTTTCATCTCTTTTAAATGCTGTGTTTCCATCAAACTTTGTTCCAGACTTATAT 2954  
Qy 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 2955 TTTCTAAACACCTGGAGACAAACACATATAATACAAACCAAGTCTGCTCTTCTCAAAATCT 3014  
Qy 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 3015 GTGAGTCAGATATCAGTGATCTTATTAGCTTTTTCACAGCCAGACATTAATGGTGACG 3074  
Qy 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaLeuAsnValMet 960  
Db 3075 ATGATTAAATGACAGTGACTATGATATCCGTGGCTCCCATAGTGGGCTTTAAATGTGTG 3134  
Qy 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 3135 CATTCAGAAAGAGCTATGTTTTTGCAGCTGTTTTCACAGTACTATGTTTATCTTTA 3194  
Qy 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 3195 CCTATATTAGTGAATATCATTAGTAACACTATCTTTATCATTTAAATGTGACTGAAACC 3254  
Qy 1001 IleGlnIleTyrSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 3255 ATCCAGATCTGGAGTACCCCATCTCTTCAAGAAATTAAGTATAGTATAGTTTTTAAATTTGAG 3314  
Qy 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 3315 CTGTATTTTCAGCAGCTTTGTTGGATCATTTGTTACTTGCAATGCCACCTTACTTTGGCC 3374  
Qy 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3375 ATGGAAATGCAGAGAAATCAATGAATCAAAAGCTTAYACTCAACTTAACTTTTCAGGTCTT 3434  
Qy 1061 LeuProSerAlaTyrTyrIleGlnAlaValValAspIleProLeuPhePheIleIle 1080  
Db 3435 TTGCCATCTGCATATTGGATTGGACAAAGCTGTTGTTGATATCCCTTATTTTATCATTT 3494  
Qy 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3495 CTATTATTTGATCTAGGAAGCTTATTGGCATTTTCATTTATGGATTATATTTTATACGTGA 3554  
Qy 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120

Db 3555 AAGTTCCTGCTGCTGTTTTTTTCCCTTATTTGGTTATGTTCCATCAGTATTCTGTCTACT 3614  
Qy 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTyrSerPheIle 1140  
Db 3615 TATATTGCTTCTTTCCACCTTTAAGAAAAATTTAAATACCAAAAGATTTTGGTCAATTATC 3674  
Qy 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3675 TATCTGTGGCAGCGTTGGCTTGATTTGATTCATCATCTGAAATAACTTTCTTTATGGGATAC 3734  
Qy 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 3735 ACAATTGCAACTATTCTTCATTATGCTTTTGTATCATCATTCATTCATTCATCTCTA 3794  
Qy 1181 GlyCysLeuIleSerPheIleLysIleSerTyrLysAsnValArgLysAsnValAspThr 1200  
Db 3795 GGTGCTGATTTCTTTTATAAAGATTTCTTGGAGAAATGTAGAAAAAATGTGGACACC 3854  
Qy 1201 TyrAsnProTyrAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 3855 TATAATCCATGGATAGGCTTTTCAGTAGCTGTTTATATCGCTTACCTGAGTGTGTACTG 3914  
Qy 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 3915 TGGATTTTCTCTTACAATACTATGAGAAAAAATATGAGGCGCAGATCAATAAGAAAGAT 3974  
Qy 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 3975 CCCTTTTTCAGAAACCTTTCAACGAAGTCTAAAAATAGGAGCTTCCAGAACCCAGAC 4034  
Qy 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 4035 AATGAGAGTGAAGATGAAGATGTCAAAGCTGAAAGCTAAAGCTAAAGTCAARAGCTGATGGGT 4094  
Qy 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 4095 TGCCAGTGTGTGAGAGAAACCATCCATTTATGTTGTCAGCAATTTGCATAAAGATATGAT 4154  
Qy 1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320  
Db 4155 GACAAAGAAAGATTTCTTCTTTCAGAAAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 4214  
Qy 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
Db 4215 TTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGGTCCAAATGGTGTCTGGCAAAAGC 4274  
Qy 1341 ThrIleLeuAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 4275 ACAATTATTAAATATTCTGTTGTTGATATTGAACCAACTTCAGCCAGGTAATTTTATGGA 4334  
Qy 1361 AspTyrSerSerGluThrSerGluAspAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 4335 GATTATTTCTCAGAGCAAGTGAAGATGATGATTTCACTGAAGTGTATGGGTACTGTCTCT 4394  
Qy 1381 GlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
Db 4395 CAGATAAACCTTTGTGGCCAGATACTCATTTGCGAGAAACATTTTGAAATTTATGGAGCT 4454  
Qy 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db 4455 GTCAAAGGAATGAGTCCAAAGTGAAGAAAGTCAATGAAGTGAATAACACATGACATCT 4514  
Qy 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
Db 4515 GATTTAAAGAAACATCTTTCAGAGAGCTGTAAGAAACTACCTGCAGGAATCAAAACGAAAG 4574  
Qy 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
Db 4575 TTGTGTTTGTCTAAAGTATGCTTAGGAAATCTCTCAGATTACTTTGTAGATGAACCATCT 4634  
Qy 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLys 1480  
Db 4635 ACAGGTATGGATCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAACATTCATTTAA 4694

QY 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500  
 Db 4695 AACAGAAAGCGGCTGCTATTCTGACCACTCATATATGGAGGCGAGAGCTGCTGT 4754  
 QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
 Db 4755 GATCAGTAGCTATCATGGTGCTCTGGGCGATTAGATGTATCGGAACAGTACAACTCTA 4814  
 QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTyrIleGluAsn 1540  
 Db 4815 AAGAGTAATTTGGAAAAGGCTACTTTTGGNAAATTAATTTGAAGGCTGGATAGNAAAC 4874  
 QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
 Db 4875 CTAGAAGTAGACCGGCTTCAAGAGAAATTCAGTATATTTTCCCAAAATGCAAGCGCTAG 4934  
 QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
 Db 4935 GAAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTCA 4994  
 QY 1581 GlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600  
 Db 4995 CAATCTTTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCAATTTGAAGATATAGCTTT 5054  
 QY 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
 Db 5055 TCTCAAGCAACATTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGAT 5114  
 QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyrPheGluArgThrGlnGluAspArgVal 1640  
 Db 5115 AATAGTTGTGGAACTTTAAACAGCACACTTTGGTGGAAACGAAACAAGAAGATAGATA 5174  
 QY 1641 ValPhe 1642  
 Db 5175 GTATTT 5180

RESULT 8  
 AX537472  
 LOCUS  
 DEFINITION Sequence 3 from Patent WO2070690.  
 ACCESSION AX537472  
 VERSION AX537472.1 GI:25269279

KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
 AUTHORS Chen, H., Kilinski, L. and le Bihan, S.  
 TITLE Abca5 transporter and uses thereof  
 JOURNAL Patent: WO 02070690-A 3 12-SEP-2002;  
 Active Pass Pharmaceuticals, Inc. (CA)  
 Location/Qualifiers

FEATURES  
 source  
 1..4917  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 4917  
 Score: 8309.00 Matches: 1629  
 Percent Similarity: 98.85% Conservative: 1  
 Best Local Similarity: 98.79% Mismatches: 2  
 Query Match: 18 Indels: 18  
 DB: 6 Gaps: 2

US-10-090-458-5 (1-1642) x AX537472 (1-4917)

QY 1 MetSerThrAlaIleArgGluValGlyValTyrArgGlnThrArgThrLeuLeuLys 20  
 Db 1 ATGTCCACTGCAATTAGGAGGTAGAGTTTGGAGACAGACCAAGCAACTTCTACTGAAG 60

QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGlnLeuLeuPheProLeu 40  
 Db 61 AATTACTTAAATTAATGCAGAACCAAAAGAGTAGTGTTCAGAAAATCTTTTCCACTA 120  
 QY 41 PhePheLeuPheTyrPheLeuIleLeuSerMetMetHisProAsnLysLysTyrGluGlu 60  
 Db 121 TTTTCTTAAATTTGGTAAATATTAATAGCATGATGCATCCAAATAAGAAAATATGAAGA 180  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80  
 Db 181 GTGCCTAATATAGAACTCAATCCTATGGACAAAGTTTACTCTTCTAATCTAATCTTGA 240  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 Db 241 TATACTCCAGTGACTAATATTAACAGCAGCATCATGCAGAAAGTGCTACTGATCATCTA 300  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 Db 301 CCTGATGTATTAATTAATCTGAAGAATATACAAATGAAAAGAAATGTTAAATCACTCCAGTCTC 360  
 QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
 Db 361 TCTAAGCCGAGCAACTTTGTAGTGGTGTTCCTCAAGAGCTCCATGTCCTATGAACTTCGT 420  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 Db 421 TTTTCTCTGATATGATTCAGTATCTTCTAATATATGATTCAGATCAAGAGCTGGCTCTCA 480  
 QY 161 LysSerCysGluAlaIleGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 Db 481 AAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTCACAGTTCCTAAGCATCCATA 540  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGluSerThr 200  
 Db 541 GATGCTGCCATATATACAGTTGAGACCAATGTTCTCTTTGGAGAGAGCTGGAGTCACT 600  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
 Db 601 AAAGCTGTTATTATGGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGATA 660  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 Db 661 ATTTTAAATATACCTAGTTATAGCATTTTCCCTTTGGATATCTTTTGGCAATTCATATC 720  
 QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 Db 721 GTAGCAGAAAAAGAAAAATAAAAGAAATTTTAAAGATAATGGGACTTCATGTACT 780  
 QY 261 AlaPheTyrPheSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 Db 781 GCCTTTTGGCTTTCCCTGGGTCTCTATATACAGTTTAAATTTTCTTATGTCCTCTCT 840  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
 Db 841 ATGGCAGTCAATGCGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTATATT 900  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 Db 901 CTGCTTTTCTCTTTATGGATTATCATCTGATTTTCTTAAATGCTGACCTCTCT 960  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 Db 961 TTTAAAAAATCAAAACATGCGGAGATAGTTGAATTTTGTACTGTGGCTTTGGATT 1020  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 Db 1021 ATTGGCTTATGATAATCTCTCATAGAAAGTTTCCCAATCGTTAGTGTGGCTTTTCAGT 1080  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 Db 1081 CCTTTCTGTGTCAGTACTTTTGTGATTGGTATGGACAGGTCAATGATTTAGAAATTTT 1140



QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
DB 1141 AATGAAGGTGCTTCATTTTCAAAATTGACTGCAGGCCCATATCTCTAATATTACAAATT 1200  
QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
DB 1201 ATCATGCTCACACTTAATAGTATATTCATGCTCTTGGCTGTCTATCTTTGATCAAGTC 1260  
QY 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLeuProSerTyrTrp 440  
DB 1261 ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTGG 1320  
QY 441 SerLysSerLysArgAsnTyrGlnLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
DB 1321 TCAAGAGCAAGAAAGAAATATGAGAGATTATCAGAGGCAATGTAAATGGAATATTAGT 1380  
QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
DB 1381 TTATAGTAATATTATGAGCCAGTTTCTTCAGAAATTTGTAGAAAGAACCCATPAGAATT 1440  
QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
DB 1441 AGTGTATATTCAGAGACATACAGAAAGAGGGTGAAATATGTGAGGCTTTGAGAAATTTG 1500  
QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
DB 1501 TCATTTGACATATATGAGGGTCAGATTACTGCCCTTCTTGGCCACAGTGGAAACAGGAAG 1560  
QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
DB 1561 AGTACATTGATGAATATCTTTGTGGACTCTGCCACCTCTTGATGGTTTGCATCTATA 1620  
QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
DB 1621 TATGGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAATGATTGGCAATT 1680  
QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
DB 1681 TGTCCACAGTTAGATATACACTTTGTGTTTGTGACAGTAGAAGAAATTTATCAATTTG 1740  
QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnValGlnLysValLeuLeu 600  
DB 1741 GCTTCAATCAAGAGGATACCAGCAACAATATAATACAGAAAGTGCAGAGGTTTACTA 1800  
QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
DB 1801 GATTTAGACATGCAGACTATCAAGATATACCAAGTAAATTAATTAAGTGTGTGTCAAAA 1860  
QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
DB 1861 AGAAAGCTGTCTATTAGGAATGTCTTCTGGAAACCAAGATACCTGCTGATGAA 1920  
QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyrArg 660  
DB 1921 CCAACAGCTGGATGGACCCCTGTTCTCGACATATTGTATGGAAATCTTTTAAATACAGA 1980  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
DB 1981 AAAGCAATCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATCTTGCA 2040  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
DB 2041 GATAGGAAGCTGTGATATACAGGAATGCTGAATGTGTGGTTCTTCAATGTTCCTC 2100  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 2101 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTATACATAGCAAAATATTGTGCCACA 2160  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
DB 2161 GAATCTCTTCTTCACATGGTAAACAAATATACCTGGAGTACTTTATTACAAACAGAAAT 2220  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760

DB 2221 GACCACACACTGTGTATAGCTTGCCTTTCAGGACATGACAAATTTTCAGGTTTGT 2280  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 2281 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCAATTCCTTATGGTGTTCATGACGACT 2340  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
DB 2341 TTGGAGAGAGTATTTTAAAGCTTAGAAGTTGAAGCAGAAATTCACCAACAGATTATAGT 2400  
QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 2401 GTATTTTACTCAGCAGCCACTGAGGAGAAATGGAATTCAAATCTTTTCATCAATGGAA 2460  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
DB 2461 CAGAGCTTACTTATCTTCTGAAACCAAGGCTGCTTAGTGAGCACCAAGCCCTTGG 2520  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
DB 2521 AAACAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTTACCTTGAACGTTGAAGTAA 2580  
QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAATTTTTTTCACAGTTTCAGATTTTATGTTT 2640  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
DB 2641 TTGGTTTCATCACTCTTTTAAATGCTGTGGTTCCTCAACAACTTGTGCCAGCTTATAT 2700  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
DB 2701 TTTTCAAAACCTGGAGACAAACCAATATAACAAACAAAGTCTGCTTCTTCAAAATCT 2760  
QY 921 Ala-----AspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
DB 2761 GTGTGTGAGAGTGTWNGTGAAGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGC 2820  
QY 935 GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
DB 2821 CAGAAACATAATGCTGACGATGATTAATGACAGTACTATGATCGTGGCTCCCATAGT 2880  
QY 955 AlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSer 974  
DB 2881 GCGGCTTTAAATGTGATGCAATTCAGAAAGGACTATGTTTTTGCAGCTGTTTTCAACAGT 2940  
QY 975 ThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHis 994  
DB 2941 ACTATGGTTTATCTTTTACCTATATTAGTGAATATCAATTAGTAACTACTATCTTTATCAT 3000  
QY 995 LeuAsnValThrGluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAsp 1014  
DB 3001 TTAATATGTGACTGAACCATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTAAGTAT 3060  
QY 1015 IleValPheLysIleGluLeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAl 1034  
DB 3061 ATAGTTTTTAAATTTAGCTGTATTTCACAGAGCTTTGCTTGGAAATCAATGTAACTGC 3120  
QY 1034 aMetProProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrG 1054  
DB 3121 AATGCCACCTTACTCTGCCATGGAAATGCGAGAAATCATATA----- 3162  
QY 1054 nLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIle 1074  
DB 3163 -----GGTCTTTTGGCATCTGATATTTGGATTGGACNAGCTGTGTGTGATAT 3209  
QY 1074 eProLeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrG 1094  
DB 3210 CCCCTTATTTTATCATCTTATTTTATGCTAGGAAGCTTATTGGCATTTTCATTTATGG 3269  
QY 1094 yLeuTyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValPr 1114

Db 3270 ATTATATTTTATACGTAAAGTTCCTTGCTGGTGGTTTTTGGCCTTATTGGTATGTTCC 3329  
 QY 1114 sSerValIleLeuPheThrTyriLeAlaSerPheThrPheLysLysLysLysLeuAenThrLy 1134  
 Db 3330 ATCACTATTCCTGTTCACTTATATTCCTTCACTTTTAAAGAAATTTTAAATACCAA 3389  
 QY 1134 sGluPheTrpSerPheIleTySerValAlaAlaLeuAlaCysIleAlaIleThrGluI 1154  
 Db 3390 AGAATTTTGGTCATTATCTATCTGTGGCAGCGTTGGCTTGTATTCGCAATCACTGAAT 3449  
 QY 1154 eThrPhePheMetGlyTyThrIleAlaThrIleLeuHisTyriAlaPheCysIleIleI 1174  
 Db 3450 AACTTTCTTTATGGGATACACAATTCGAATTCCTTCAATATGCTTTGGTATCATCAT 3509  
 QY 1174 sProIleTyProLeuLeuGlyCysLeuIleSerPheIleLysLysLysLysSerTrpLysAsnVa 1194  
 Db 3510 TCCAATCATCCACTTCAGTTCAGTTCCTTCTTCAATAAGATTTCTTGGAAAGATGT 3569  
 QY 1194 lArgLysAsnValAspThrTyAsnProTrpAspArgLeuSerValAlaValIleSerPr 1214  
 Db 3570 ACGAAAAATGTGGACACCTATATCCATGGATAGGCTTTCAGTAGCTGTATATCGCC 3629  
 QY 1214 oTyriLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyTyriGluLysLysTyriGly 1234  
 Db 3630 TTACCTGCAGTGTGACTGTGGATTTTCTCTTCAATACATCTATGAGAAAAATATGAGG 3689  
 QY 1234 lArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLy 1254  
 Db 3690 CAGATCAATAGAAAGATCCCTTTTCAGAACTTTCAGAACTTCAAGAGCTTAAATAGAA 3749  
 QY 1254 sLeuProGluProAspAsnGluAspGluAspGluAspValLysAlaGluArgLeuLy 1274  
 Db 3750 GCTTCCAGAACCCAGACCAATAGAGATGAAGATGAAGATGCAAGCTGAAGACTAAA 3809  
 QY 1274 sValLysGluLeuMetGlyCysGlnCysGlnGluLysProSerIleMetValSerAs 1294  
 Db 3810 GGTCAAGAGCTGATGGGTSCCAGTGTGTGGAGGAAACCATTCATTCGTACGCAA 3869  
 QY 1294 nLeuHisLysGluTyAspAspLysLysAspPheLeuLeuSerArgLysValLysVa 1314  
 Db 3870 TTTTGATAAAGATATGATGACAAAGAAAGATTTCTTCTTTCAAGAAAGTAAAGAAAGT 3929  
 QY 1314 lAlaThrLysTyriLeSerPheCysValLysGlyLysGlyLysLeuGlyLeuGlyPr 1334  
 Db 3930 GGCACATAATACATCTTTCTGTGTGAAAAGGAGAGATCTTAGGACTATTGGGTCC 3989  
 QY 1334 oAsnGlyAlaGlyLysSerThrIleAlaAsnIleLeuValGlyAspIleGluProThrSe 1354  
 Db 3990 AAATGGTGTGCAAAAGCACAAATATTAATATTCGTGTGGTGATATGAACCAACTTC 4049  
 QY 1354 rGlyGlnValPheLeuGlyAspTySerSerGluThrSerGluAspAspAspSerLeuLy 1374  
 Db 4050 AGCCAGATATTTTATGAGATATTTCTTCAGAGCAAGTGAAGATGATGATTCACCTGAA 4109  
 QY 1374 sCysMetGlyTyCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHi 1394  
 Db 4110 GTGATGGGTACTGCTCCTCAGATAAACCCCTTGTGGCAGATACATTCATTCGAGGAACA 4169  
 QY 1394 sPheGluIleTyGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSe 1414  
 Db 4170 TTTTGAAATTTATGAGCTGTCAAGAGGATGAGTGCAGATGACATGAAGAAGTCATAAG 4229  
 QY 1414 lArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuPr 1434  
 Db 4230 TCGAATAACACATGACCTGATTTTAAAGAACATCTTCAGAGAGCTGTAAAGAAATACC 4289  
 QY 1434 oAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleTh 1454  
 Db 4290 TCGAGGAATCAAAACGAAAGTTGTGTTTGGCTCTAGTATGTAGGGAATCCTCAGATATC 4349  
 QY 1454 rLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAl 1474  
 Db 4350 TTTTGCTAGATGATACCATCTACAGGTATGGATCCCAAGCCAAACAGACATGTGGCGGC 4409

QY 1474 aIleArgThrAlaPheLysAsnArgLysArgAlaIleLeuThrThrHisTyMetGl 1494  
 Db 4410 AATTCGAACGTGATTTAAACAGAAAGCGGCTCTATTCAGCACTCACTATATGGA 4469  
 QY 1494 uGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysI 1514  
 Db 4470 GGAGGCAGAGGCTGTCTGTGATCTGATGATCATGGTGTCTGGCAGTTAAGATGAT 4529  
 QY 1514 eGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyriPheLeuGluIleLysLe 1534  
 Db 4530 CGGAACAGTACACATCTAAAGAGTAAATTTGGAAAAAGCTACTTTTGGAAATTTAAAT 4589  
 QY 1534 uLysAspTrpIleGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyriLePh 1554  
 Db 4590 GAAGGCTGGATAGAAACCTAGAGTAGTACCGCTTCAAGAGAGAAATTCAGTATATTT 4649  
 QY 1554 eProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAlaTyriLysIleProLysGl 1574  
 Db 4650 CCAAAATGCAAGCGCTCAGAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAAGGA 4709  
 QY 1574 uAspValGlnSerLeuSerGlnSerPhePheLysLeuGluAlaLysHisAlaPheAl 1594  
 Db 4710 AGATGTTCACTCCCTTTCACATCTTTTTHAAGCTGGAGAGCTTAAACATGCTTTTC 4769  
 QY 1594 aIleGluGluTyriSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLy 1614  
 Db 4770 CATTTGAAGAAATATAGCTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAA 4829  
 QY 1614 sGluGlnGluGluGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluAr 1634  
 Db 4830 AGAACAGAGGAGGAGAGATATAGTTGTGGAACCTTAAACAGCACACTTTGGTGGGAACG 4889  
 QY 1634 gThrGlnGluAspArgValValPhe 1642  
 Db 4890 AACACAGAGATAGATAGTAGTATTT 4914  
 RESULT 9  
 AX537470  
 LOCUS AX537470  
 DEFINITION Sequence 1 from Patent WO02070690.  
 ACCESSION AX537470  
 VERSION AX537470.1 GI:25269277  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1  
 AUTHORS Chen, H., Kilinski, L. and le Bihan, S.  
 TITLE Abc5 transporter and uses thereof  
 JOURNAL Patent: WO 02070690-A 1 12-SEP-2002;  
 Active Pass Pharmaceuticals, Inc. (CA)  
 FEATURES  
 source  
 1. 5463  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 5463  
 Score: 8309.00 Matches: 1629  
 Percent Similarity: 98.85% Conservative: 1  
 Best Local Similarity: 98.79% Mismatches: 2  
 Query Match: 98.61% Indels: 18  
 DB: 6 Gaps: 2  
 US-10-090-458-5 (1-1642) x AX537470 (1-5463)  
 QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
 Db 114 ATGTCCACTGCAATTAGGAGGAGTAGGAGTTTGGAGACAGACAGAACACTTCTACTGAAG 173



2334 GACCAACAACTTGTGTATAGCTTGCCTTCAAGGACATGGCAAAATTTTCAGGTTTGGTT 2393  
QY SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
2394 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCATTTCTTATGGTGTTCCTCAGCAGCT 2453  
QY LeuGluAspValPheLeuLeuLeuValGluAlaGluIleAspGlnAlaAspTyrSer 800  
2454 TTGGAAGACGTAATTTTAAAGCTAGAGTTGAAGCAGAAATTCACCAAGCAGATATAGT 2513  
QY ValPheThrGlnGlnProLeuGlnGluGluMetAspSerLysSerPheAspGluMetGlu 820  
2514 GTATTACTCAGACGCCCTGAGAGAAATGGATTCAAAATCTTTTGGATGAATGGAA 2573  
QY GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
2574 CAGAGCTTACTTATCTTCTGAACCAAGGCTGCTTAGTGAGCACCATCAGCCTTGG 2633  
QY LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
2634 AAACCAACAGATATACATAGCAAAAGTTTCATTTCTTTACCTGAAACGTTGAAGTAA 2693  
QY SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
2694 TCAGTGAGATCAGTGTGCTTCTGCTTCTTAAATTTTTCACAGTTCAGATTTTATGTT 2753  
QY LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
2754 TTGGTTTCATCACTCTTTTAAAAATGCTGGTTCCTCCATCAAACTTGTTCAGACTTAT 2813  
QY PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGluAsnSer 920  
2814 TTTCATAAACCTGGGACAAACCATATATACAAACCAAGTCTGCTTCTTCAAAATCT 2873  
QY Ala-----AspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
2874 GCTGGTGAGAGTGTTNNGTCAAGACTCAGATATCAGTGATCTTATTAGCTTTTCAAGC 2933  
QY GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
2934 CAGAACATTAATGGTGACGATGATTATGACAGTGACTATGATCCGTGGCTCCCATAGT 2993  
QY AlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSer 974  
2994 CGGGCTTTAAATGTGATGATTCAGAAAGGACTATGTTTTTGGAGCTGTTTTCAACAGT 3053  
QY ThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHis 994  
3054 ACTATGGTTTATCTTTACCTATATAGTGAATATCATAGTAACATCTATCTTTATCAT 3113  
QY LeuAsnValThrGluThrIleGlnIleTyrSerThrProPhePheGlnGluIleThrAsp 1014  
3114 TTAAATGTGCTGAAACCATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCAT 3173  
QY IleValPheLysIleGluLeuTyrPheGlnAlaAlaLeuLeuGly-IleIleValThrAl 1034  
3174 ATAGTTTTTAAATTTAGAGCTGATTTTCAAGCAGCTTTGCTTGGAAATCATTTAACTGC 3233  
QY aMetProProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrG1 1054  
3234 ATGCCACCCCTACTCTGCCATCGAAATTCAGAGATCATAA----- 3275  
QY nLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspI1 1074  
3276 -----GGTCTTTTGGCATCTGCATATTTGGATTGGACAGCTGTTTGTATAT 3322  
QY eProLeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrG1 1094  
3323 CCCCTTATTTTATCACTTCTTATTTTGTATGTAGGAGCTTATTGGCATTTTCATTATGG 3382  
QY YLeuTyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValPr 1114

3383 ATTATATTTTATTACTGTAAAGTTCTTGCTGTGGTTTTTTTTCCTTATTGGTTATGTTC 3442  
QY oSerValIleLeuPheThrTyrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrIy 1134  
3443 ATCAGATTATCTGTTCACTTATATTGCTTCTTTTACCTTTAAGAAAATTTTAAATACCA 3502  
QY sGluPheTyrSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluI1 1154  
3503 AGAAATTTTGGTTCATTTATCTATCTGTGCGACGTTGGCTTGTATTGCAATCAGTGAAT 3562  
QY eThrPhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIle11 1174  
3563 AACTTCTTTTATGGGATACAAATTCCAACTATTCTTCAATTATGCTTTTGTATCATCAT 3622  
QY eProIleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTyrPlyAsnVa 1194  
3623 TCCAACTATCCACTTCTAGGTGCTGATTTCTTTCATAAAGATTTCTTGAAGAATGT 3682  
QY lArgLysAsnValAspThrTyrAsnProTyrAspArgLeuSerValAlaValIleSerPr 1214  
3683 ACAGAAAATGTGGACACCTTATATCCATGGGATAGGCTTTTCAGTAGCTTGTATATCGC 3742  
QY oTyrLeuGlnCysValLeuTyrIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyG1 1234  
3743 TTACCTGCGAGTGTACTGTGGATTTTCTTACAATACTATGAGAAAAAATATGAGG 3802  
QY YArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLy 1254  
3803 CAGATCAATAAGAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTTAAATAAGAA 3862  
QY sLeuProGluProProAspAsnGluAspGluAspValLysAlaGluArgLeuLy 1274  
3863 GCTTCAGAACACACACAGACAAATGAGGATGAAGATGAAGTGTCAAAGCTGAAAGACTAA 3922  
QY sValLysGluLeuMetGlyCysGlnCysGlnGluLysProSerIleMetValSerAs 1294  
3923 GGTCAAAGAGCTGATCGGTGCGAGTGTGTGAGGAGAAACCATCCATTATGTCAGCAA 3982  
QY nLeuHisLysGluTyrAspAspLysAspPheLeuLeuLeuSerArgLysValLysLysVa 1314  
3983 TTTGCTATAAGATATGATGACAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAGT 4042  
QY lAlaThrLysTyrIleSerPheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyPr 1334  
4043 GGCACCTAAATACATCTTCTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTTGGTTC 4102  
QY oAsnGlyAlaGlyLysSerThrIleIleAsnIleLeuValGlyAspIleGluProThrSe 1354  
4103 AAATGGTGTGGCAAAAGCACAAATTATTAATTCTTGGTGGTGTATTTGAACCACTTC 4162  
QY rGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAspAspSerLeuLy 1374  
4163 AGGCCAGGTATTTTATAGGAGATTTATCTTCAGAGACAAAGTGAAGATGATGATCACTGA 4222  
QY sCysMetGlyTyrCysProGlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHi 1394  
4223 GTGTATGGGTACTGCTTCAGATAAACCTTTGTGGCCAGATACTACATTGCGAGGAACA 4282  
QY sPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSe 1414  
4283 TTTTGAATTTATGAGCTGTCAAAGGAATGAGTGCAGTCATGAAAGAGTCTATAAG 4342  
QY rArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuPr 1434  
4343 TCGAATAACACATGCTGATTTTAAAGAACATCTTTCAGAGACTGTGAAAGAACTACC 4402  
QY oAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleTh 1454  
4403 TGCAGGAATCAACAGAAAGTGTGTTTGTCTTAAGTATGCTAGGGAATCTCTCAGATTAC 4462  
QY rLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAl 1474  
4463 TTGCTAGATGAACCACTCTACAGTATGGATCCCAAGGCCAACACACACATGTGGGAGC 4522

QY 1474 aileArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeuThrHisTyrMetG1 1494  
 Db 4523 AATTCGAAGTGCATTTAAACACAGAAACGGGCTGCTATTCTGACCACTCACTATATGGA 4582  
 QY 1494 uGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysI1 1514  
 Db 4583 GGAGCGACAGGCTGCTCTGTATCGAGTAGCTATCATGTGTCTGGGCGAGTTAAGATGAT 4642  
 QY 1514 eGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLe 1534  
 Db 4643 CGGACAGTACACACTAAAGAGTAATTTGGAAAGGCTACTTTTGGAAATTAAT 4702  
 QY 1534 uLysAspTrpIleGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePh 1554  
 Db 4703 GAAGACTGGATAGAAACCTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTT 4762  
 QY 1554 eProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAlaTyrLysIleProLysG1 1574  
 Db 4763 CCCAATGCAAGCGTCAAGAAAGTTTCTTCTATTTTGGCTTAAATTCCTAAGGA 4822  
 QY 1574 uAspValGlnSerLeuSerGlnSerPhePheLysLeuGluAlaLysHisAlaPheA1 1594  
 Db 4823 AGATGTCAGTCCCTTTCCACATCTTTTAAAGCTGGAAGACTAAACATGCTTTTGC 4882  
 QY 1594 aileGluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLy 1614  
 Db 4883 CATGAAGAATATAGCTTTCTCAAGCAACATTGGAACAGTTTTGTAGAACTCACTAA 4942  
 QY 1614 sGluGlnGluGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluAr 1634  
 Db 4943 AGAACAGAGGAGGAAGATAATAGTTGTGGAACITTTAAACAGCACACTTTGGTGGGAACG 5002  
 QY 1634 gThrGlnGluAspArgValValPhe 1642  
 Db 5003 AACCAAGAAGATAGTAGTATTT 5027

RESULT 10  
 AX417826 4785 bp DNA linear PAT 18-JUN-2002  
 LOCUS AX417826  
 DEFINITION Sequence 3 from Patent WO0231147.  
 ACCESSION AX417826  
 VERSION AX417826.1 GI:21522943  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE  
 1 Hu,Y. and Nepomnichy,B.  
 AUTHORS Human transporter proteins and polynucleotides encoding the same  
 TITLE Patent: WO 0231147-A 3 18-APR-2002;  
 JOURNAL LEXICON GENETICS INC (US)  
 FEATURES  
 Location/Qualifiers  
 1..4785  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 4785  
 Score: 8112.00 Matches: 1583  
 Percent Similarity: 99.69% Conservative: 0  
 Best Local Similarity: 99.69% Mismatches: 5  
 Query Match: 96.27% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-090-458-5 (1-1642) x AX417826 (1-4785)  
 QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
 Db 1 ATGTCCACTGCAATTAGGAGGAGTAGGAGTTTGGAGACAGACAGACAGACACTTCTACTGAAG 60

QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGluIleLeuPheProLeu 40  
 Db 61 AATTACTTAAATTAATGCAGAACCAAAAGAGTAGTGTTCAGAAATTTCTTTTCCACTA 120  
 QY 41 PhePheLeuPheTrpIleuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
 Db 121 TTTTATTTTATTTGGTAAATTAATAGCATGATGCATCCAAATAAGAAATATGAAGAA 180  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
 Db 181 GTGCTTAATATAGAACTCAATCTATGACAGATTTACTCTTTCTTAATCTAATTTCTTGA 240  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 Db 241 TATACTCCAGTGAATAATTAAGAGCATCATGCAGAAAGTGTCTACTGATCATCTA 300  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 Db 301 CTTGATGTCATTAATTAATGAGAAATATACAAATGAAGAATAATGTTAATACATCCAGTCTC 360  
 QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
 Db 361 TCTAAGCCGAGCAACTTTGTAGGTGTGGTTCATAAGACTCCATGTCTCTATGAACCTCGT 420  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 Db 421 TTTTTCCTGATATGATCCAGTATCTTCTATTAATGGAATCAAGAGCTGGCTGTTC 480  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 Db 481 AAATCATGTGAGGCTGCTCAGTACTGCTCAGGTTTCACAGTTTACAGCATCCATA 540  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
 Db 541 GATGCTGCCATATATACAGTTGAAGACCAATGTTCTCTTTGGAAGAGCTGGAGTCAACT 600  
 QY 201 LysAlaValIleMetGlyGluThrAlaValAlaGluIleAspThrPheProArgGlyVal 220  
 Db 601 AAAGCTGTTATATGGAGAAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGATA 660  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 Db 661 ATTTTAATATACCTAGTATATAGCATTTTCCACTTTTGGATACTTTTGGCAATTCATATC 720  
 QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 Db 721 GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGACTTCATGATACT 780  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 Db 781 GCCTTTTGGCTTTCCTGGGTCTCTCTATAWACAAGTTTAAATTTTCTTAATGTGCTCTCT 840  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
 Db 841 ATGGCAGTCAATTCGACAGCTCTTTGTATTTCTCTCAAGTAGCAGCATTTGTGATATT 900  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 Db 901 CTGCTTTTTCCTTTATGGAATTAATCATCTGTATTTTGTCTTAAATGCTGCTACACTCTT 960  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 Db 961 TTTTAAAAATCAAAACATGTGGAAATAGTTGAAATTTTGTGTGCTGGCTTTTGGATTT 1020  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 Db 1021 ATTTGGCTTATGATAATCTCATAGAAAGTTTCCAAATCGTTAGTGTGGCTTTTCAGT 1080  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 Db 1081 CCTTCTGCTACTGACTTTTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTT 1140

QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleThrIle 400  
DB 1141 AATGAGGTGTTTCAATTTCAATTTGACTGCAGGCCCATATCTCTAAATATTACGATT 1200  
QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
DB 1201 ATCATGCTCACACTTAATAGTATATCTATGCTCTTGGCTGTCTATCTTGATCAAGTC 1260  
QY 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTyr 440  
DB 1261 ATTCCAGGGGAATTGGCTTACGAGATCATCTTTATATTTCTGAAGCCTTCATATTGG 1320  
QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
DB 1321 TCAGAGCAAAAGAAATATGAGGAGTTATCAGAGGCAATGTTAAATGGAAATATTAGT 1380  
QY 461 PheSerGluIleLeuGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
DB 1381 TTTAGTGAATTTTGAAGGAGTTCTTCAGAAATTTGTAGGAAAGAGGCCATAAGAAAT 1440  
QY 481 SerGlyIleGlnLysThrTyrArgLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
DB 1441 AGTGATTTACAGACACATACAGAAAGAGGCTGAAATATGAGAGGCTTTGAGAAATTTG 1500  
QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
DB 1501 TCATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCCACAGTGGACAGGAAG 1560  
QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
DB 1561 AGTACATTTGATGAATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCACTATA 1620  
QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
DB 1621 TATGACACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATTGGGCATT 1680  
QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
DB 1681 TGTCACAGTTAGATATACACTTTGATGTTTGACAGTAGAAGAAATTTATCAATTTTG 1740  
QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
DB 1741 GCTTCAATCAAGGATACCCAGCCCAATATAATACAGAGAGTGCAAGAGTTTACTA 1800  
QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
DB 1801 GATTTAGACATGCAGACTATCAAGATTAACAGCTTAAATTAAGTGGTGGTCAAAA 1860  
QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
DB 1861 AGAAAGCTGTCTTAGGAATTTGCTTCTTGGGAACCCAAAGATCTGCTCTAGATGAA 1920  
QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLysTyrArg 660  
DB 1921 CCACAGCTGGAAATGGAACCCCTGTTCTGCACATATTTGATGGAATCTTTAAATACAGA 1980  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
DB 1981 AAAGCCAAATCGGTGACAGTTCAGTACTCATTTTCATGGATGAAGCTGACATCTTGCA 2040  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
DB 2041 GATAGAAAGCTGTGATATCACAGGAATGCTGGAATGTTGGTTCTTCAATGTTCTCTC 2100  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 2101 AAAAGTAAATGGGGATCGGTACCGCTGAGCATGTACATAGACAAATATTGTGCCACA 2160  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
DB 2161 GAATCTCTTTCTCTGCTGGTAAACACATATACCTGGAGCTACTTTATTACACAGAA 2220  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760

DB 2221 GACCAACAACCTGTGTATAGCTTGCCTTTCAAGGACATGGACAAATTTTCAGGTTGT 2280  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 2281 TCTGCCCTAGACAGTCATTCAAATTTGGTGGCATTTCTTATGGGTTTCCATGACGACT 2340  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGluAlaAspTyrSer 800  
DB 2341 TTGGAAGACGTTATTTTAAAGCTAGAGTTGAAGCAGAAATTCACCAAGCAGATTATAGT 2400  
QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 2401 GTATTTACTCACAGCCACTGGAGGAGAAATGGATTCAAATCTTTTGTGAATGGA 2460  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
DB 2461 CAGAGCTTACTTATCTTTCTGAAACCAAGGCTKCTCTAGTGAGCACCATGAGCCTTTGG 2520  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
DB 2521 AAACACACAGATGATACAAATAGCAAGTTTCAATTTCTTACCTTGAAACGTGAAGTAA 2580  
QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 2581 TCAGTGAGATCAGTGTGCTCTCTTAAATTTTTCACAGTTTCAGATTTTATGTTT 2640  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProLysLeuValProAspLeuTyr 900  
DB 2641 TTGGTTCATCCTCTTTTAAATGCTGGTTCCTCAAACTTTGTTCCAGACTTATAT 2700  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
DB 2701 TTTCTTAAACCTGGAGACAAACACATAATAACAAACAAAGTCTGCTCTTCAAATTTCT 2760  
QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
DB 2761 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAAGCCAGAACATAATGGTGAC 2820  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
DB 2821 ATGATTAATGACAGTGACTATGATCGTGGCTGCCCATAGTGGCTTTAAATGTGRTG 2880  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
DB 2881 CATTCAGAAAGACATATGTTTTCAGCTGTTTTCACACAGTACTATGTTTATTTCTTTA 2940  
QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
DB 2941 CCTATTATTAGTGAATATCATTAGTAACACTATCTTTATCATTTAAATGTGCTGAACC 3000  
QY 1001 IleGlnIleTyrSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
DB 3001 ATCCAGACTGGAGTACCCCATTTCTTGAATCATTTGTTACTGCAATGCCACCTTACTTGGC 3120  
QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAla 1040  
DB 3061 CTGTATTTTCAGCACCTTTGCTTGAATCATTTGTTACTGCAATGCCACCTTACTTGGC 3120  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
DB 3121 ATGGAAATGTCAGAGAATCATAGATCAAAAGCTTAYACTCAACTTAACTTTTCAGTCTT 3180  
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
DB 3181 TTGCCATCTGCATTTGGATTGGACAGCTGTGTTGATATCCCTTATTTTATCATTT 3240  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
DB 3241 CTTATTTTGTGCTAGGAGCTTATTTGGCATTTTCAATATGATGATATATTTTATCTGTA 3300  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120







Db 1656 TATGGACACAGAGCTCTCTGAATAGATGAATGTTTGAAGCGAGGAATGATTCGCATA 1715  
QY CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
Db 1716 TGTCCACAGTCAGATAGAACTTTGATGTTCTGACAGTAGAAGAAAATTTATCAATTTTG 1775  
QY AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
Db 1776 GCTTTCAGTCAAGAGATACCGCCACCAACATTAATTCAGAGAGTACAGAAAGTTTACTG 1835  
QY AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys 620  
Db 1836 GATTTGCGACATGCAAGCCATCAAGATAATCAAGCGAAAAAACTAAGTGTGGCCCAAAAG 1895  
QY ArgLysLeuSerLeuGlyIleAlaValLeuGlyValAsnProLysIleLeuLeuAspGlu 640  
Db 1896 AGAAAGCTGCTTTAGAAATTCAGATTCCTGGAAATCCAAAGATCCTCTCTGTAGATGAG 1955  
QY ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
Db 1956 CCGACAGCAGGAATGACCCCTCTCGCCACATTTGTAATCTTCTCAAGTACAGG 2015  
QY LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 2016 AAGGCTAAGCGAGTGACAGCTCTCAGCACTCACTTCATGGATGAGCGCGACATTCCTGCC 2075  
QY AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2076 GACAGGAAGCTGTATATACACAGGAATGCTGMAATGTTGGTTCTTCAATTTCTCTG 2135  
QY LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2136 AAAAGTAATGGGGAATCGGCTACCGCTCTGAGCATGTATATAGACAGGTACTGTGCCACA 2195  
QY GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db 2196 GAGTCGCTCTCTGCTGCTGTGTAGACAGCACATCCCTGCGCGCTCTGCTGACAGAGAT 2255  
QY AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 2256 GACCAGCAGATGCTCTATAGCTTGCCCTTCAAAGACATGGACAAATTTTCAGGCTGTGT 2315  
QY SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2316 TCTGCTCTAGACATTCATTCAAACCTGGGTGTAAATTCCTATGGTGTTCATAGACGACA 2375  
QY LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2376 TTGGAAAGCGTGTTTTAAAGCTAGAAGTTGAAGCAGAAATTCACCAAGCAGATATAGT 2435  
QY ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2436 GTATTTACCCAGCAGCCGAGGAGAAACAGATTCAAAGTCTTTTGTGAAATGGAA 2495  
QY GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 2496 CAGAGTTTACTTATTCTCTGAAACCAAAGCTTCTTAGTCAGCACCATGAGTCTCTGG 2555  
QY LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
Db 2556 AAGCAGCAAGTACTCAATAGCAAGTTTCATTTCTCTACTGAAAGGAGAAACGAAA 2615  
QY SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2616 TCAGTCGCTCTGTGTTGCTTCTGCTTTAAATTTTGTGAGTTCAGATTTTATGTT 2675  
QY LeuValHisHisSerPheLysAsnAlaValProIleLysLeuValProAspLeuTyr 900  
Db 2676 TTGGTCCATCATTTCTTTAAAAATGCTGTGGTTCCCATCAAGCTCGTTCAGACTGTAT 2735  
QY PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920

Db 2736 TTTCTAAACCTGGAGATAAACCTCATAAATACAAACAAGCCCTGCTGCTTCAAAATCT 2795  
QY AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 2796 ACTGATTCAGATATCATGATCTTATTGACTTTTTCACAGCAGACATAAATTTGTGGCC 2855  
QY MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 2856 ATGTTTAAATGACAGTGACTATGTCGTCGCTCTCATAGTGGGCTCTCAATGTGGTG 2915  
QY HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 2916 CAGTCAGAAAAGGACTATGTTTACTGCTGTTTTCACAGTAGTACTATGTTTATCTTCTG 2975  
QY ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 2976 CCAGTAATGATGAACATCATTAGTACTACTATCTATCATTTAAATGATGACTGACACT 3035  
QY IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 3036 ATCCAGATCTGGAGTACCCCGTTTCAATCAGAAATTAATGACATTGTTTAAAGTTGAG 3095  
QY LeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 3096 CTATATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCC 3155  
QY MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3156 ATGGAAATCGAGAGATCATAGATCAAGCTTACACTCAGCTTAAACTTTTCAGGCTTT 3215  
QY LeuProSerAlaTyrTrpIleGlyGlnAlaValAspIleProLeuPhePheIleIle 1080  
Db 3216 TTGCCATCCGCCCTACTGATGGATTGGACAGGCTGTGGTGGACATCCCTGTTTGTGTGT 3275  
QY LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3276 CTGACTTTGATGCTGGGAAGTTTATTGCTCTTTCATCAGGAGCTGATTTTATCTCTGTA 3335  
QY LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 3336 AAGTTCTTCTGCTGGTGTGTTTTCCTCATTTGCTTATGTGCCCTCGGTTATCTGTTCACT 3395  
QY TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 3396 TATATAGCTTCGTTTACCTTTTAGAAAATTTTAAATACCAAGGAATTTGCTCATTTATC 3455  
QY TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3456 TATTCTGTGACAGCAATTGGCTTGTGTGGCAGTCAGCGAAATAAATCTTCTTCTGGGGTAT 3515  
QY ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 3516 GGAGTTACGGCTGTTTTCATTTACCTTCTGATAGCCATCCCAATCTACCCGCTCTG 3575  
QY GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
Db 3576 GGTTGTCTGATTTCTTCTATAAAGGTTCTTGGAGAATATATACAAAACCTGAGAAATGCC 3635  
QY TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 3636 TACAATCCCTGGGACAGACTTTTAGTTGCTGTAAATCATGCCCTACCTCGAGTGTGACTG 3695  
QY TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 3696 TGGATTTTCTCTTACAACTATGAGAAAAACATGGAGGCAGATCAATAAGAAAGGAT 3755  
QY ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 3756 CCGCTTTTCAGGGCCCTCTCACAAAAGCCAAATGAAGGTTTCCAGAACCCACCATC 3815  
QY AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 3816 AATGAGGATGAGATGAAGATGTCAAAGCGAAAAGTTGAAGTTTAAAGACTGATGGGT 3875



Qy	1	MetSerThrAlaIleArgGluValGlyValTyrArgGlnThrArgThrLeuLeuLeuLys	20
Db	9	ATGGCTACTGCAATTAGGGATGTGGCGCTTTGGAGACAGACCAGAACTCTTCTCTGAAG	68
Qy	21	AsnTyrIleuIleLysCysArgThrIlyssSerSerValGlnGluIleuPheProLeu	40
Db	69	AATTACCTCGCAANTGAGACTAAAGAGTAGTGTTCAGGAATTCCTTTTCTCTCTA	128
Qy	41	PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu	60
Db	129	TTTTTCTTATTTTGGCTGATATTAATTAGCATGATGCATCCAAATAGAAATATGAGAA	188
Qy	61	ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleuGly	80
Db	189	GTATCTGATATGAACTCAGCCCTATGACAAATCCATCCTTTCCAACTTATCTCGGGG	248
Qy	81	TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu	100
Db	249	TACACTCCAGTGACTAATACTAACAGCAGCGTCAATCAGAGAGTGTTCACCGATCATCTT	308
Qy	101	ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu	120
Db	309	CCGACGTTCTAGTTTACTGGAAGATACCCCATGGAAGAGAGCTGCTGGCATCCAGTTTG	368
Qy	121	SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg	140
Db	369	TCTAAGCCACGAACTTCGTGGGTGGTTCCTCAAGACGTCATGCTCATGAATCGGG	428
Qy	141	PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer	160
Db	429	TTTTTCTGATATGGTTCCAGTGTCCTCTGTTATATGGATTCAAGACGAGCGTGTTC	488
Qy	161	LysSerCysGluAlaAlaAsnLysTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle	180
Db	489	AAGTCATGTGATGTGCTCAGTACTGGTCTTCAGGGTTTACAGCTTTACAGCCCTCAATA	548
Qy	181	AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGluSerThr	200
Db	549	GATGCTGCCATATATACAGTTCAAGACCAATGTTCTCTGTGGAGAGCTGGAGTCAACC	608
Qy	201	LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal	220
Db	609	AAAGCTGTCAATCATGGGGAGCTGCGGTCTGGAGATTGACCTTCCACAGAGGGTCT	668
Qy	221	IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle	240
Db	669	ATCCTCATCTACCTCGTCATAGCCTTCTCGCCCTTCGGGTACTTCTGGCAATTACATT	728
Qy	241	ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr	260
Db	729	GTACGAGAAAGGAAAGAGGTTAAAGAGTTTTTAAGATAATGGGACTTCACGACACT	788
Qy	261	AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu	280
Db	789	GCTTTTGGCTGTCTCGGGTCTCTCTGTACACGAGTTTGATCTTCTTATGTCTGCTTCT	848
Qy	281	MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe	300
Db	849	ATGGCCGTCAATGTCACAGCGTCTCGTTATTCCTTCAGAGTAGCAGCATCGTGATTTTT	908
Qy	301	LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu	320
Db	909	CTGCTCTCTTCTTCGTGCGGATTGTCATCTGTGTTTTTGTCTTAAATGTTGACGCCCT	968
Qy	321	PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe	340
Db	969	TTTTAAAAATCAAGACAGTGGGAGTCGTTGAGTTTTTTGTACCGTGGTGTGGCTTT	1028
Qy	341	IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer	360
Db	1029	GTGGCTGTTCATAGTCTCGTAGAAGATTCCCAAGTTCACAGTCCGTAGTGTGCTCTT	1088

Qy	361	ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe	380
Db	1089	CCITTTGTTGTCAGTGGCGCTTTCTCTGATTGGGATTGCAcAGGTcATGcATTAGAAgATTTC	1148
Qy	381	AsnGluGlyAlaSerPheSerAnLeuThrAlaGlyProTyrProLeuIleIleThrIle	400
Db	1149	AATGAAGAGCGCTTATTCTCTAGTTTGACTGAAGGCCCTcACCGCTTAAATTATCACCCCTT	1208
Qy	401	IleMetLeuThrLeuAnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal	420
Db	1209	ACCATGCTCGCTGGACAGTGTTCTATGCCCTGCTGGCTGTGTACCTTGATcAGGTC	1268
Qy	421	IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp	440
Db	1269	ATCCAGGGGAATTTGGCTTTGAGAGGTcATCTTTGTACTTCTTAAAGCCATCATATGG	1328
Qy	441	SerLysSerLysArgAsnTyrGluGluLeuSerGlyGluAsnValAsnGlyAsnIleSer	460
Db	1329	TCAAAAACAAGGAACATATAAGAGCTATCAGAGGGCAACATCAATGGGATATTAGT	1388
Qy	461	PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle	480
Db	1389	CTCAATCAAAATTTGTAGCCCGTTCTTCTCAGAATTTATAGAAAGAAAGTATAAGAAATA	1448
Qy	481	SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu	500
Db	1449	AGTGGTATTcAGAAAGCCTATAGAGAAGAAATGAGACCGTGAGGGCTTTTGAGAAATTTG	1508
Qy	501	SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys	520
Db	1509	TCGTTTGACATCTATGAAGCCAGATTACTGCTCTGCTGGCCcACAGTGAACAGGAAAA	1568
Qy	521	SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerProSerAspGlyPheAlaSerIle	540
Db	1569	AGCACATAATGAATATTCTGTGTGACTGTGCCcACTTCTGATGGTTGTCTCTATATA	1628
Qy	541	TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle	560
Db	1629	TATGGACACAGATCTCTGAAATAGATGAATGTTTGAAGCGAGGAAAAATGATTGGCATA	1688
Qy	561	CysProGlnLeuAspIleHisPheAspValLeuThrValGluAlaAsnLeuSerIleLeu	580
Db	1689	TGTCCACAGTCAGATATGAACCTTTGATGTTCTGACAGTAGAAGAAATTTTCAATTTTG	1748
Qy	581	AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu	600
Db	1749	GCITTCAGTCAAGGAATACCAGCCACACACATPAATTCAGGAAGTACAGAAAGTTTACTG	1808
Qy	601	AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys	620
Db	1809	GAITTTGACATGCAAGCCATCAAGAATATCAAGCGAAAAAACTAAAGTGGTGCCAAAAAG	1868
Qy	621	ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu	640
Db	1869	AGAAAGCTGCTTTAGGAATTCAGTCTCTTGGGAATCCAAAGATCCTCTGCTAGATGAG	1928
Qy	641	ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg	660
Db	1929	CCGACAGCAGGAATGACCCCTCTCTCGCCcATTTGTTGGAAATCTCTCAAGTACAGG	1988
Qy	661	LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla	680
Db	1989	AAGGCTAACCGAGTGACGTCCTTCAGACTTCATTCGATGAGGCCcACATCTTTGCC	2048
Qy	681	AspArgIlyAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu	700
Db	2049	GACAGGAAAGCTCATATCAcAGGAATGCTGAAATGTTGTTGTTCTTCAATTTTCTCTG	2108
Qy	701	LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr	720
Db	2109	AAAAGTAATATGGGGAATCGGTACCGTCTGAGCATGTATATAGACAGGTACTGTGCCACA	2168
Qy	721	GluSerLeuSerSerLeuValLysGlnHisIleProGlyValaThrLeuLeuGlnAsn	740



```
QY 1455 uLeuaspGluProSerThrGlyMetAspProlysalalysGlnHisMetTrpArgAlaI 1475
D 4389 GGTGATGAGCGCTGTACAGGTATGAGCCCAACAGCACATGTGGCAGGCAT 4448
QY 1475 eArgThrAlaPheLysAsnArgAlaAlaIleLeuThrHisTyrMetGluG 1495
D 4449 TCGAACTGGTTTAAACACAGAGCGGCGCTCTCTACGCGATTAACAGGA 4508
QY 1495 uAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleG 1515
D 4509 GCGGAGGCGTGTCTGTGACAGGCGCCCATCATGTGTCCGCGCAGCTAAGATGCA 4568
QY 1515 yThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluLysLeuL 1535
D 4569 ACAGTACACATCTAAAGAGAGCTTTGGCAAGCGCTACITTTTAGAAATTAACATA 4628
QY 1535 sAspTrpIleGluAenLeuGluValAspArgLeuGlnArgGluLysGlnTyrIlePhePr 1555
D 4629 GCAGTGGATAGAAAACCTGGAATAGATCGCTTCAAGAGAAATTCATATATTTCCC 4688
QY 1555 eAlaAlaSerArgGlnGlnSerPheSerSerIleLeuAlaTyrLysIleProLysGluAs 1575
D 4689 AAATCAAGCCGCCCAAGAGAGCTTCTCTTATCTGCTTATAAAATCCAAAGAGGA 4748
QY 1575 pValGlnSerLeuSerGlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaI 1595
D 4749 TGTCAGTCTCTCTACAGCTCTTTTGTAGTGGAGGAGCAAAACACACATTTTGGCAT 4808
QY 1595 eGluLysTyrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysG 1615
D 4809 TGAAGAAATACAGCTTTCTCAAGCAACCTCGAAGAGCTTTTGTAGAACTTACTAAGGA 4868
QY 1615 uGlnGluGluGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTrpTppGluArgTh 1635
D 4869 GCAGGAGGAGGAGATATAGCTGGCGGCTTTAAACAGCAGCTCTGCTGGGGAAGAG 4928
QY 1635 rGlnGluAspArgVal 1640
D 4929 GCAGGAAGACAGAGTC 4944

RESULT 13
AB097675
LOCUS AB097675 4929 bp mRNA linear ROD 01-APR-2003
DEFINITION Mus musculus abca5 mRNA for ABC transporter subfamily A mABCA5,
complete cds.
ACCESSION AB097675
VERSION AB097675.1 GI:29420876
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Kubo, Y., Satsuma, Y., Sekiya, S., Nada, S. and Yamaguchi, A.
TITLE Molecular cloning of mABCA5, the mouse homologue of ABCA5
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4929)
AUTHORS Kubo, Y., Satsuma, Y., Sekiya, S., Nada, S. and Yamaguchi, A.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2002) Yoshiyuki Kubo, I.S.I.R., Osaka Univ.;
Mihogaoka 8-1, Ibaraki City, Osaka 567-0047, Japan
(E-mail: kubo37@anken.osaka-u.ac.jp, Tel:81-06-6879-8546,
Fax:81-06-6879-8549)
FEATURES
source
1..4929
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="11"
/map="11E1"
/tissue_type="brain"
/dev_stage="newborn"
```

## gene

```
1..4929
/genes="abca5"
1..4929
/genes="abca5"
/codon_start=1
/product="ABC transporter subfamily A mABCA5"
/protein_id="BAC66658.1"
/db_xref="GI:29420877"
/translation="NATAIRDVGVMRQTRTLKLLKYLKICRTKKSGVQILPFLPLF
WILVSMHPNKKEEYEDSLSPMDKFSLSNVILGYTPTNITSSIMORVTDHLPLK
VTVTEYANKEKLVAASLKSNGFVWVFDKTMSEYELFFPMIPVSSYVMSREBGS
KTCDAQYNSLQGTVLQASIDAAIQLKTNVMSGELESTKAVIMGEAAVVIDTPPR
GVLLIVTAFSPFGVFLAHTVAEKKLKEFLKIMGLHDTAFWLSWLLVSLIFEL
MSLLMAVITASSLFPQSSIVIFLLFLYGLSSVFFPALMLTFLPKSKHGVGVZEV
TVVFGVGLIIVLIESFPRSLVWLFSPCCAPLIGIAGVMELEDFNEGALFNSITEG
PPLIITIIIMLADSVVVLAVLDQVTPGFGFLRRSLYFLKPSYWSKNKNTKEL
SEGNINGNISLNEIVEPVSEFIEKEAIRISGIQKSYRKTENVALRNLSDIYEGQ
ITALLHGSGTGLTNILGLCPSPDGFASVYGHVSEIDEMFEAKMIGICPQSDI
NFDVLTVBENLSILASIKGIPANNIIQEVKLLDLDMOAIKONAKKLSGQKRLS
VGLAVLGNPKILLDEPTAGMDPCSRHIVNLLKYRKANRVTVFSTHFMDEADILADR
KAVISQMLKCVGSSIFLKSQWGIYVLSWIDRYCATESLSLYRQHLPAALLOON
DQQLYSLPFKMDKFSGLFSLDLSNLSVSYGVSMITLEDVFLKLEVEAIEDQAD
YSVFTQCPREETDSKSFDEMSLILSETKASSVSTNSLMKQVSTIAKTHFLSLK
RESKSVRAVLLLLLFFAVQIEMFLLHSFKNAVPIKLVPLDYFLKPDGPKHKLTS
LLQNSTDSINGLIIEFFAHQNMVAMFENDSYVSAAPHSAALNVRSKDYVFSAVF
NLTQWVCLPVMMNIIISNYLYHLNVTAEIQWSTPFIQIETIDIVFKIELYFOALLGI
IVTAMPYFAMENAEHNKIKAYTKLSGLSPAYMVGQAVVDIPLFFVLLMLGSL
PAPHGLYFPAKELAVNFCILAYVPSVILFYIASPTFKILNTKEWFSYVYAL
ACVATETTFLOQYAVTAVFYHFCIAPIYPLGLCLISPIKSGWKNPKNENTNPW
DELLVAVIMPYLCILWIPLOHYEKHGSRIRKDPFRALSKAKNKKFPEPPINE
DEEDVPAERLAKVXELMGCCCEERPAIMVNLNKHEDDKDFLRKNTTKVATYIS
FCVKKGEILGLGPNAGKSTVINTLVGDEPTEGSKIPLGIDYSHSDEDESIKMGY
CQTNPLPDLTLOHFETYGAVKGMSPGDKVEVSRITKALDLKHELQKTVKPLPAG
IKKLCFPLSMVGNPQVLLDDEPSTGMDPRAKQHMWRARITAFKNKKAALLTHYME
EAEKCDVAIVMNSQLCIGTVQHLKSKFGKGYFLKLDKWIENLETRDQRIQY
IFPNSARQSESPSSILAFKIPKEDVQSLQSPAKLEBAKRTFAIEYVSPQAILEQVFV
ELTKQSESDNCGSLASTLWERTQEDRVVF"
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 0 Length: 4929
Score: 7729.00 Matches: 1480
Percent Similarity: 94.88% Conservative: 78
Best Local Similarity: 90.13% Mismatches: 84
Query Match: 91.73% Indels: 0
DB: 10 Gaps: 0
US-10-090-458-5 (1-1642) x AB097675 (1-4929)
QY 1 MetSerThrAlaIleArgGluValTyrArgGlnThrArgThrLeuLeuLys 20
D 1 ATGGCTACTGCAATAGGATGGAGTGGAGAGTTGGAGACAGACCAGAACACTTCTACTGAAA 60
QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGluIleLeuPheProLeu 40
D 61 AATTACTTAATTAATGCAGGACTTAAATAAGTAGTGTTCAGAAAATCTTTTCTCTTA 120
QY 41 PhePheLeuPheTrpIleLeuLeuLeuSerMetMetHisProAsnLysLysTyrGluGlu 60
D 121 TTTTCTTATTATTTGGCTGATATTAGTCATGATGTCATCCAAATAAGAAAATATGAAGAG 180
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80
D 181 GTATCTGATATAGAGCTCAGCCCTATGGACAAATTCAGCCTTTCACAGCTTATCTTGA 240
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
D 241 TACACTCCCGTGACTAACATTACAGCAGCATTAAGCAGAGGGTTTCTACCATCATCTTT 300
QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
D 301 CCCAGGTTTATAGTTACTGTAAGAAATACGCAATAGAAAGAACTGGTAGCCCAAGTCTT 360
```

QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
 DB 361 TCTAAGTCCAGCAACTCGTAGTGTGGTTTCAAGACACCATGTCATAGAACTTCGT 420  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 DB 421 TTTTTCCTGAAATGATCCAGTGTCTTCTATTTATATGAATCAAGAGAGGCTGTCA 480  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 DB 481 AAGACATGTGATGCTGCTCAGTACTGCTTGGGGTTTACAGTCTCGAGGCATCGATA 540  
 QY 181 AspAlaAlaIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGlnLysThr 200  
 DB 541 GATGCTGCCATTATACAGCTGAAGACCAATGTTCTGTGTGAGCGAGCTGGAGCTCGACC 600  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
 DB 601 AAAGCTGTGATCATGGGAGAGCCCTGTTGTGGAGATTGACCTTCCCGCGAGGGGTC 660  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 DB 661 ATCCTCATCTACCTCGTCATAGCCTTCGCGCCCTCGGCTACTTCTGGCCTATCCATC 720  
 QY 241 ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 DB 721 GTGGCAGAAAAGAAAGAAAGTAAAGAAATTTTAAAGATAATGGGACTTCATGACACT 780  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 DB 781 GCTTTTGGCTTCTCGGCTCTCTGACGCAAGCTTGATTTTCTTATGCTCCCTGCTT 840  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300  
 DB 841 ATGGCTGTATCCCAACAGCTCTCTGTTATTCCTCAGAGTAGCAGCATTTGATCTTT 900  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 DB 901 CTACTGTCTCTTATATGGATGTGTCATCTGTGTTTTTGTCTTAATGCTGACGCCCTTT 960  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 DB 961 TTTAAAAAATCAAAACACGTGGAGCTCGTTGAGTTTTTGTCCACGTGGTGTGGATTT 1020  
 QY 341 IleGlyLeuMetIleIleLeuIleLysSerPheProLysSerLeuValTrpLeuPheSer 360  
 DB 1021 GTGGCCCTGCTGATTTGCTCATAGAAAGTTTCCCGAGCTCGCTGGTGTGCTCTT 1080  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 DB 1081 CTTTGTGTGCTGAGTGTGCTTCTGATTTGGATTGGACAGGTGATGATTTAGAGATTTT 1140  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 DB 1141 AATGAGGGTGCCTTATTTCTAATTTGACTGAAGGTCCCTATCTCTAATTTACTATT 1200  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 DB 1201 ATCATGCTAGCTCTTGACAGTGTTCTATGTCCTCTGCTGGCTGTGATCTCGACCAAGTC 1260  
 QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 DB 1261 ATTCCAGGGGAATTTGGCTTGAGGAGGTGATCTTTGTATTTTGAAGCCATCGTATTGG 1320  
 QY 441 SerLysSerLysArgAsnTyrGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 DB 1321 TCAAAAAACAAGAAACATATAGGAGCTATCGAGGCAACATTAATGGCAATATTAGT 1380  
 QY 461 PheSerGluIleIleLeuProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
 DB 1381 CTCATGAATTTGTTGAGCCCGTTCTTCTCAGAAATTTATAGGAAAGAGCTATAGAATA 1440  
 QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyLeuAsnValGluAlaLeuArgAsnLeu 500

DB 1441 AGTGGTATTCAAGAAATCTATAGAAAGAAACTGAGACGTGGAGCTTTGAGAAATTTG 1500  
 QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 DB 1501 TCAATTTGACATATATGAAGTCAAGTACTGCTGCTGGGCGACAGTGGAAACAGAGAAA 1560  
 QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
 DB 1561 AGCACACTGATGAATATTCTGTGTGACGTGTCTCACCCCTCTGATGGGTTTCTCTATA 1620  
 QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 DB 1621 TATGGACACAGAGTCTCTGAAATAGATGAATGTTTGAAGCAAGAAATAATGATGGCATA 1680  
 QY 561 CysProGlnLeuAspIleHisPheAspValIleThrValGluGluAsnLeuSerIleLeu 580  
 DB 1681 TGTCCGACGTAGATATAAATCTTGTATGTTCTGACAGTAGAAGAAATTTATCAATTTG 1740  
 QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 DB 1741 GCTTCAATCAAAGGAATACAGCAACATAATAATTCAGAAAGTGCAGAAAGTGTGCTG 1800  
 QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys 620  
 DB 1801 GATCTGGACATGCAAGCCATCAAGATATCAAGCGAAAGATTAAGCGGTGGTCAGAAA 1860  
 QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
 DB 1861 AGGAAGCTCTGTAGGAATTCAGTTCCTCGGAATCCAAAGATATCTCTGCTAGACGAG 1920  
 QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyrArg 660  
 DB 1921 CCTACAGCAGGAATGGACCCCTGCTCTGCCATATTTGTTGGAATCTTCAAGATATAGA 1980  
 QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 DB 1981 AAGGCTAACAGAGTGACCGTGTGTAGTACTCACTTCATGGATGAGGCTGACATTTCTGCC 2040  
 QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 DB 2041 GACAGAAAGCTGTCTATATCAAGAAATGCTGAAAGTGTGTTGTTCTTCAATTTCTCTA 2100  
 QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 DB 2101 AAGATTAATGGGAATCGGCTACCGCTGACATGATATATAGACAGATCTGTGCCACA 2160  
 QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrIleLeuGlnAsn 740  
 DB 2161 GAGTCACTGTGCTGCTGTAGGACGACATACCCGACGCGCTACTGACAGCAGAAAT 2220  
 QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
 DB 2221 GACACAGCTCGTGTACAGCTGCTTTCAAAGCATGGACAAATTTTCAGGCTGTGTT 2280  
 QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
 DB 2281 TCTGCTCTAGACATTCAATCAAACCTGGGTGTTATTTCTTATGGTCTTTCCATGACACA 2340  
 QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
 DB 2341 TTGGAAGATGTAATCTTAAAGCTTAGAAGTTGAAGCAGAAATTTGACCAACAGCATATAGT 2400  
 QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
 DB 2401 GTATTTACACAGCAGCGCGGGGAGGAGAAACAGATTTCAAAATCTTTTCAATGAAATGGAA 2460  
 QY 821 GlnSerLeuLeuIleLeuSerGlnThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
 DB 2461 CAGAGTTTACTTATCTCTCTGAAGCAAGAGTCTTCTCAGTGAGCACCACATGAGCCTCTGG 2520  
 QY 841 LysGlnGluMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860



Db 2521 AAGCAGCAAGTGTCTACGATTGCAAGTTTCATTTCTCTCTCATTTGAAACGAGAAAGCAAA 2580  
Qy 861 SerValArgSerValLeuLeuLeuLeuLeuLeuPhePheThrValGlnIlePheMetPhe 880  
Db 2581 TCAGTGGCGCTGTGTCTCTCTGCTTTTAAATTTTTCAGTTCAGATTTTATGTTT 2640  
Qy 881 LeuValHisHisSerPheLeuAsnAlaValValProIleLeuLeuValProAspLeuTyr 900  
Db 2641 TTTCTCCATCAATCTTTTAAANAATGCTGTGTTCCCATCAAACTCGTTCAGACTGTAT 2700  
Qy 901 PheLeuLysProGlyAspLysProHisGlySerLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 2701 TTCCTAAAGCTGGAGATAAACCCTCATAAAATACAAAACAAAGCCTGCTCTCAAAATCT 2760  
Qy 921 AlaAspSerAspIleSerAspLeuLeuSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 2761 ACTGACTCAGATATCAATAGTGTCTTATTAGATTTTTCACACAGAACATAAATGGTGGCA 2820  
Qy 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 2821 ATGTTTAATGACAGTGACTATGTCTGCTCTCTCACAGTGGGCTCTCAATGTGGTG 2880  
Qy 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 2881 CGTTCGAAAGGACTATGTTTTCCTGCTGTTTCAACAGTACTATGGTTTATTGTTG 2940  
Qy 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 2941 CCAGTCATGATGACATCATAGTAATCTATCTTATCATTAATATGACTGAGCC 3000  
Qy 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 3001 ATCCAGACCTGGAGTACCCCTTCATTCAGAAATTAATGACATTTGTTTAAATTTGAG 3060  
Qy 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 3061 CTATATTTTCAAGCAGCTTTCCTTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCC 3120  
Qy 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3121 ATGGAAATGAGAGAAATCATAGATCAAGCTTATATCTCAACTTAAACTTTCGGGTCTT 3180  
Qy 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
Db 3181 TTGCCATCAGCCTATTTGGTTGGCAAGCTGTTGTGATATTCCTTTGTTTGTGTT 3240  
Qy 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3241 CTGATTTTGTGCTGGGAAGTTTATTTGTCATTTTCATCATGAGCTGTATTTTATCCTGCA 3300  
Qy 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 3301 AAAATTTCTGTGTGGTGTGCTGCTCATTTATGTGCTTATGTGCTCCCGTCATTTCTTCACG 3360  
Qy 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 3361 TACATAGCTTCTGTTCACTTCAAGAAATTTTAAATACCAAGGAATTTTGTGTCATTTATC 3420  
Qy 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3421 TATTCTGTGACCGCATTTGGCTGTGTGGCAATCATCGAAACAACTTTCTTCTGCAATAT 3480  
Qy 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 3481 GCAGTTACGGTGTCTTTCATACCTTCTGCAATAGCCATTCATCTACCTCTCTCTCTG 3540  
Qy 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
Db 3541 GGTGTCTGATTTCTTTCATAAAGGGTCTTGGGAATATATGCCAAAATATGAGAAATACC 3600  
Qy 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 3601 TACAAATCCCTGGGATAGACTTTTAGTTGCTGTAAATATGCGCTTACCTCCCTGATGTA 3660

Qy 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyLysArgSerIleArgLysAsp 1240  
Db 3661 TGGATTTTCTCTTCAACACTATGAGAAATATACATGGAGGCAGATCTATATAGAAGGAT 3720  
Qy 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 3721 CCATTTTTCAGGGCCCTTTCACAAAAGGCCAAAATAAGAGTTTCCAGAACACCCATC 3780  
Qy 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 3781 AATGAGGATGAAGATCAAGATCTCAAGCTGAAGACTGAAGCTGAAGCTGAAGCTGATGGT 3840  
Qy 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 3841 TGCCAGTGTGTGAGGAGAGCCAGCCATTTATGTTATGCAATTTGCTAAGAAGATGATGAT 3900  
Qy 1301 AspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer 1320  
Db 3901 GACAAAGAAAGATTTTCTCATTTCAAGAAAAACAAACAAAGTAGCAACAAATATCATCTCT 3960  
Qy 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValAclyLysSer 1340  
Db 3961 TTCTGTGTGAAAAAGGAGAGATCTTGGACTGTGTGGGTCCAAATGGAGCTGCAAAAGC 4020  
Qy 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 4021 ACAGTCATTAACACTCTGTTGGGACGTTGAGCCAACCTTCAGGCAAGATCTTTCTAGGA 4080  
Qy 1361 AspTyrSerSerGluThrSerGluAspAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 4081 GATTATGTTTCATTTCAAGTGAAGATCAGCAATCCATTAAGTGTATGGGTACTGCCCC 4140  
Qy 1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
Db 4141 CAGACAAACCCACTCTGGCCAGATCTCACTCTACAGGAAACATTTTGAGATTTACGAGCT 4200  
Qy 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db 4201 GTGAAAGGAATGAGTCCCTGGTGACATGAAGGAAGTCATCAGTCGAATAACAAAACCTTT 4260  
Qy 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
Db 4261 GATTTGAAAGAAACATCTTCAGAAAGCTGTAAAGAAGCTACTCTGAGGATCNAGGGAAG 4320  
Qy 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
Db 4321 CTGTGTTTGTCTCTCAGCATGCTGGGAAACCTCAGTGCAGCTGCTGATGAGCCGTCC 4380  
Qy 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLys 1480  
Db 4381 ACAGGATGGACCCCAAGAGCCAAAGCAGCATGTGGAGAGCTATTCGAACTGCATTTAAA 4440  
Qy 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500  
Db 4441 AACAAAAAGCGGCGCCCTCTCTGACTACTCATTTACATGGAAGAGGAGAGCTGCTGT 4500  
Qy 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
Db 4501 GACAGAGTGGCCATCATGTTATCTGGCAGCTTAAGATGTTTGGAAACAGTACCAATCTTA 4560  
Qy 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
Db 4561 AAGAGTAAAGTTTGTGTAAGGCTACTTTTGGAAATTAATTAAGGACTGGATAGAAAAC 4620  
Qy 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
Db 4621 CTGGAAAAATAGATCGCCTTCAAGAGAAATCAATATATTTTCCCAAAATGCAAGCCGCA 4680  
Qy 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
Db 4681 GAGAGCTTTCTCTTCTATTTTGGCTTTTAAATTCACAAAGGAAGATGTTTCAGTCTCTCTCA 4740

QY 1581 GlnSerPheLeuGluGluAlaLysHisAlaPheAlaLeuGluGluTyrSerPhe 1600  
 Db 4741 CAATCTTCGCTAGTTGGAAGAGCAACGCACTTTGCAATGGAATACAGCTTT 4800  
 QY 1601 SerGlnAlaThrLeuGluGluValPheValGluLeuThrLysGluGluGluGluAsp 1620  
 Db 4801 TCTCAAGCAACCTCGAAGACAGCTTTTGTAGAACTTACTAAAGAGCAAGAGGAAGAT 4860  
 QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640  
 Db 4861 AACAGCTGTGGACCTTTGGCCAGCACTCTCTGTGGGAGAGGACGAGGAGGATAGAGTA 4920  
 QY 1641 ValPhe 1642  
 Db 4921 GTATTT 4926  
 RESULT 14  
 AF491842  
 LOCUS  
 DEFINITION Mus musculus ATP-binding cassette transporter sub-family A member 5  
 ACCESSION AF491842  
 VERSION AF491842.1 GI:22087247  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 5347)  
 AUTHORS Annilo, T., Chen, Z.-Q., Shulenin, S. and Dean, M.  
 TITLE Evolutionary analysis of a cluster of ATP-binding cassette (ABC)  
 Genes  
 JOURNAL Mamm. Genome 14 (1), 7-20 (2003)  
 MEDLINE 22419899  
 PUBMED 12532264  
 REFERENCE 2 (bases 1 to 5347)  
 AUTHORS Annilo, T., Chen, Z.-Q., Shulenin, S. and Dean, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-MAR-2002) Laboratory of Genomic Diversity, National  
 Cancer Institute, Frederick, MD 21702, USA  
 FEATURES  
 Location/Qualifiers  
 1..5347  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 1..5347  
 /gene="Abca5"  
 210..5338  
 /gene="Abca5"  
 /codon\_start=1  
 /product="ATP-binding cassette transporter sub-family A  
 member 5"  
 /protein\_id="AA090895.1"  
 /db\_xref="GI:22087248"  
 /translation="MATAIRDVGWVRQTRLLKNYLKCRKKSSVQBELFPLPFLF  
 WILVMHPKPKVEEDSIELSPMDKFSNLNVLGYTPVTNITSSIMQVSDHLEPK  
 VIVTBEYANKELEVAASKSNSFVGVKDTMSYELRFFPEMIPVSIYVNSREGCS  
 KTCDAOYWSLGFVLQASIDAAIIIOLETNVSVNSELESTKAVIMGEAAVVEIDTPR  
 GVLLYLVATSPRCGYFLAIHVAEKKLEFKLGLHDTAFWLSWLLYASLIPL  
 MSLNAVATASLSPQSSIVIFLLVFLYGLSSVFPALMTLPFKSKHGVVEVFP  
 TVVFGVGLLVILSEFPRSLVWLPSCQCAFLIGIAQVNHLEDNEGAFSLNLTG  
 PYPLITIMALDSVFVLLAVLDQVIGPFGRLRSLLYFLKPSYVSKNKNRYEL  
 SEGNTHGINSIEIVPVSSEFTGKEAIRISGIQSKYRKKTENVEALRNLSFDYEGQ  
 ITALLGNSITGLWNLTCGPPSPGASIVHRVSEIDEIPEAKRMIGICPQSDI  
 NFDVLTVENSLASIKGIPANNIIQEVQVLLDNOAIKDNOAKLISGQKRLSL  
 VGIVLGNPKILLDEPTAGMDPCSRHIVNLLKYRKANRVTFSTHMDRADILADR  
 KAVISQMLKCVGSSIFLKSXWIGYRLSMYIDRYCATESLSLVROHIIPAALLOON  
 DQQLVTSFQKMDKPSGLFSALDISHNIGVISGVSNITLEDVFLKLEVAEDQAD  
 RESKSVRAVILLLLIFFAVQIPMWELHESFKNAVVPKLVDPDLYFLKPKDPKPKYKTS  
 YSVFTQQPREGSDSKSDEMEQSLILSETKASSVSNPSPWQKQVSTIAKPHFLSLN  
 LLLQNSTDSINGLIEFFAHQNIWAMFNDSYVSAAPHSALANVRSEKDYVPSAVF  
 NSTWVYCLPVMNNIISNYLVYHLNVTETAIQTWSTPFIQEIETDIFVKIELYFOAALLGI

## ORIGIN

Alignment Scores:  
 Pred. No.: 0 Length: 5347  
 Score: 7674.00 Matches: 1471  
 Percent Similarity: 94.46% Conservative: 80  
 Best Local Similarity: 89.59% Mismatches: 91  
 Query Match: 91.08% Indels: 0  
 DB: 10 Gaps: 0  
 US-10-090-458-5 (1-1642) x AF491842 (1-5347)  
 QY 1 MetSerThrAlaIleAArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
 Db 210 ATGGCTACTGCAATAGGAGTGTGGAGTTTGGACAGACAGCCAGACACTTCTACTGAAA 269  
 QY 21 AsnTyrLeuIleLysCysArgThrLysSerValGlnGluLeuPheProLeu 40  
 Db 270 AATTACCTAAATTAATGAGGACTAAATAAGTAGTGTTCAGGAAATCTTTTCTCTCTA 329  
 QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
 Db 330 TTTTCTTATTTTGGCTGATATTAGTACGATGATGATCCAAATAAGAAATATGAGAG 389  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80  
 Db 390 GTATCTGATATAGAGCTCAGCCCTATGGACAAATTCAGCCTTTCACAGCTTATTCTTGA 449  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 Db 450 TACACTCCCGTGACTACATTAACAGCAGCATATGACAGGGGTTTCTACCATCATCTT 509  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 Db 510 CCCAAGGTTTATAGTTACTAGAAATACGCAATGAGAAAGAACTGTCAGCCGCAAGTCTT 569  
 QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
 Db 570 TCTAAGTCCAGCAACTTCGTAGGTGGTGTTCCTCAAGACACCATGTCCTATGAACTTCG 629  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 Db 630 TTTTTCCTGAAATGATTCAGTGCTTCTTATATATGATTAATCAAGAGAGGCTGTTC 689  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 Db 690 AAGACATGATGCTCTCTCAGTACTGTCTTTGGGGTTTACAGTTCTGAGGATCATGATA 749  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGluSerThr 200  
 Db 750 GATGCTGCCATATATACAGTGGAGACCATGTTCTGTGTGGAGGAGCTGGAGTCGACC 809  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
 Db 810 AAAGCTGTGATCATGGAGAGGCGCGTGTGTGGAGATTGACACCTTCCCGGAGGGGTC 869  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 Db 870 ATCTCATCTACTCTCTGATAGCTTCTGCGCTTCTGGCTACTTCTCTGCAATCCACATC 929  
 QY 241 ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260

Db 930 GTGGCAGAAAAAGAAAGATTAAAGAAATTTTAAAGATAATGGGACTTCATGACACT 989  
Qy 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
Db 990 GCTTTTGGCTTCTCTGGGTCTCTCTGACCAAGCTTGATTTTCTTATGTCCCTGCTT 1049  
Qy 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
Db 1050 ATGGCTGTCATCGCAACAGCTTCTTCGTTATTCCCTCAGAGTAGCAGCATTTGATCTTT 1109  
Qy 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
Db 1110 CTACTGGCTCTCTTATATGGATGTCATCTGTGTTTTTGTCTTAAATGCTGAGCCCTCTT 1169  
Qy 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
Db 1170 TTTAAAAAATCAAAACACGTGGAGTCGTTCAGTTTTTTGTCCCGTGGTGTGGATTT 1229  
Qy 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
Db 1230 GTTGGCCTGCTGATTGCTCATAGAAAGTTTCCCAAGTCCGCTGGTGGCTCTTCACT 1289  
Qy 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
Db 1290 CCTTGTGTCAGTGTGCTTCTGATGGGATTCACAGGTGATGCAATTTAGAGATTTT 1349  
Qy 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
Db 1350 AATGAGGTGCTGCTTATTTCTTAATTTGACTGAAGTCCCTATCTCTAAATTTACTATT 1409  
Qy 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
Db 1410 ATCATGCTAGCTTTGACAGTGTCTTATGCTCTCTGCTGTGTATCTCGACCAAGTC 1469  
Qy 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
Db 1470 ATTCCAGGGGAATTGGCTTGAGGAGTCATCTTTGTATTTTGAAGCCATCGTATGG 1529  
Qy 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
Db 1530 TCAAAAAACAAAGAAATATAGAGCTATCGAGGGCAACATTAATGGCAATATTAGT 1589  
Qy 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
Db 1590 CTCATGAAATGTTGAGCCGCTTCTCAGAAATTTATAGGAAAGAGCTATAAGATA 1649  
Qy 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
Db 1650 AGTGGTATTCAGAAATCCTATAGAAAGAAACTCGAAACCTGAGAGGCTTTGAGAAATTTG 1709  
Qy 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
Db 1710 TCATTTGACATATATAGAGTCAAGTACTGCACTGTGGCCACAGTGGAAACAGGAA 1769  
Qy 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540  
Db 1770 AGCACACTGATGATATTTCTGTGGACCGGTGCTCCACCTCTGATGGTTTTGCTCTATA 1829  
Qy 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db 1830 TATGGACACAGAGTCTCTGAAATAGATGAAATATTTGAAGCAAGAAATGATTTGGCATA 1889  
Qy 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
Db 1890 TGTCGGCAGTCAGATATAAATCTTTGATGTTCTGACAGTAGAAGAAATTTATCAATTTG 1949  
Qy 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnIleValGlnLysValLeuLeu 600  
Db 1950 GCTTCAATCAAGGAATACCAAGCAACATATAATTCAGAAAGTGCAGAAAGTGTGGCTG 2009  
Qy 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
Db 2010 CATCTGGACATGCAAGCATCAAGATANTCAAGCGAAAGATTAGCGGTGTGTCAGAA 2069

Qy 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
Db 2070 AGAAAGCTGTCTGAGGAATTCAGTTCTCGGAATCCAAAGATATCTCTCTAGACGAG 2129  
Qy 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyrArg 660  
Db 2130 CCTACAGCAGGAATGGACCCCTGCTCGCCATATTTGTTGGAATCTTCTAAAGATATAGA 2189  
Qy 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 2190 AAGGCTAAACAGAGTAGCCGTGTTTGTACTCTACTTCATGTAGTGGAGCTGACATTTTGGC 2249  
Qy 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2250 GACAGGAAGCTGTCATATCACAGGAATGCTGAAGTGTGTGTTCTTCAATTTCTTA 2309  
Qy 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2310 AAGAGTAAATGGGAATCGGCTACCGCTGAGCATGTATATAGACAGATCTGTGCCACA 2369  
Qy 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
Db 2370 GAGTCACTGCTGCTGTTAGGAGCAGCATACCGCAGCCGCTACTGACGAGAT 2429  
Qy 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 2430 GACCAGCAGCTGCTGTACAGCTGCTTCAAAGACATGGACAAATTTTCAGGCTTGT 2489  
Qy 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2490 TCTGCTCTAGACATTCATCAAACCTGGGTGTTATTTCTTATGGTGTTCATGACACA 2549  
Qy 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2550 TTGGAAGATGTTATCTTAAAGCTAGACAGTTGAAGCAGAAATTGACCAAGCAGATATAGT 2609  
Qy 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2610 GTATTTACACAGCAGCCCGGAGGAGGAGAAACAGATTCAAATCTTTGATGAAATGGAA 2669  
Qy 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 2670 CAGAGTTTACTTATCTCTCGAAACGAAGCTTCTTCAGTGAGCACCATGAGCCCTGG 2729  
Qy 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db 2730 AAGCAGCAAGTGTCTACGATTGCAAGGTTTCATTTCTCTCATTAACCGAGAAAGCAA 2789  
Qy 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2790 TCAGTGGCGCTGTGTTCTGCTTCTGCTTTTAAATTTTTCAGTTCAGATTTTATGGTT 2849  
Qy 881 LeuValHisHisSerPheLysAsnAlaValValProLysLysLeuValProAspLeuTyr 900  
Db 2850 TTTCTCCATCATCTTTTAAATAATGCTGTGTTCCCATCAAACTCGTTCAGACTGTAT 2909  
Qy 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
Db 2910 TTCCTAAGCCCTGGAGATAACCTCTATAAATACAAACAGCCCTGCTGCTTCAAAATCT 2969  
Qy 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 2970 ACTGACTCAGATATCAATGGTCTTATTGAGTTTTTTCACACCACCAACATAATGTGGCA 3029  
Qy 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 3030 ATGTTTAAATGACAGTACTATGTGCTGCTCTCTCAGTGGCGCTCTCAATGTGTG 3089  
Qy 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 3090 CGTTCTGAAAGGACTATGTTTTTCTGCTGTTTCAACAGTACTATGTTTATGTTTGT 3149

```

QY 981 ProfileLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000
Db .....
QY 3150 CCAGTCATGAGAACATCATAGTAACCTATCTTATCATTTAATGTGACTGAAGCC 3209
QY 1001 IlednIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020
Db ATCCAGACCTGGAGTACCCCGTTCAATCAAGAAATTAATGACATTTGTTTTAAATTTGAG 3269
QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040
Db CTATATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCC 3329
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060
Db ATGGAAATGCAGAGAATCATAGATCAAGATCAAAAGCTTATATCTCAACTTAAACTTTCCGGTCTT 3389
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080
Db TTGCATCAGCCTATTTGGTTGGACAAGCTGTTGGATATTCCTTTGTTTTTGTGTGT 3449
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100
Db CTGATTTTGTGCTGGGAAGTTTATTTGCATTTTCATCATGCACTGTATTTTATCTGCA 3509
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120
Db AAATTTCTGCTGGTGTGTTTGGCTCATTCCTTATGTGCTCCCGTCATCTTGTTCACG 3569
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140
Db TACATAGCTTCGTTCACTTTCAAGAAATTTTAAATACCAAGGAATTTTGTCTATTATC 3629
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160
Db TATTCGTGACCGCATGCTGCTTGTGCAATATCGCAATACCGAAACAACTTTCTTCTGCAATAT 3689
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180
Db CGAGTTACGGTGCTTTTCATTACCTTCTGCAATGCCATTCCAATCTACCTCTCTCTG 3749
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200
Db GGTTCCTCGATTTCTTTTCAATAAGGGTCTTTGGGAAGATATGCCAAAAATAGAGATACC 3809
QY 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220
Db TACAATCCCTGGGATAGACTTTTGTGCTGTATATCATGCCCTACCTGCAAGTGTACTG 3869
QY 1221 TrpIlePheLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240
Db TGGATTTTCTCTTACAACACTATGAGAAATATACATGGAGCAGATCTATAAGAAAGGAT 3929
QY 1241 ProphePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260
Db CCATTTTTCAGGCGCTTTTCAAAAGGCCAAAATAAGAAAGTTTCCAGAACCAACCCATC 3989
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280
Db AATGAGGATGAAGATGAAGTGTCAAGCTGAAAGACTGAAGGTATAAGAGCTGATGGGT 4049
QY 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300
Db TGCCAGTGTGTGGAGGAAGCCAGCCATATGTGTATGCAATTTGCAATAAGATATGAT 4109
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320
Db GACAAAGAAAGATTTTCTTCAATCAAGAAATAACAAACAAAGTAGCAACAAATAATCATCTCT 4169
QY 1321 PheCysValLysLysGlyGluIleLeuGlyLeuGlnGlyProAsnGlyAlaGlyLysSer 1340
Db TTCTGTGTGAAAGAGAGATCTTGGACTGTGTGGTCCAAATGGAGCTGGCAAAAGC 4229
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360

```

```

Db .....
QY 4230 ACAGTCATTAAACACTCTGTTGGGACCTTGAGGCAACTTCAGGCAAGATCTTTCTAGGA 4289
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380
Db GATTATGTTTCACTTCAAGTGAAGATGACCAATCCATTAAGTGTATGGGTACTGCCCC 4349
QY 1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400
Db CAGACAAACCCACCTGTGCCAGATCTCACTCTACAGGAACATTTTGAGATTTACGGAGCT 4409
QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420
Db GTGAAAGAATGAGTCCCTGGTGACATGAAGAAGTCATCAGTCGAATTAACAAAGCCCTT 4469
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440
Db GATTTGAAAGACATCTTCAGAGACTGTAAAGAGCTACCTGCAGGGATCAAGCGAAG 4529
QY 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460
Db CTGTGTTTGTCTCTCAGCATGCTGGGAAACCTCAGGTGAGCGTCTCTGGATGAGCGCTCC 4589
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480
Db ACAGGATGAGACCCCAAGAGCCAGCAGCAATGTGGAGAGCTATTCGAACCTGCATTTAA 4649
QY 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500
Db AAAAAAGCGGCGCGCTCTCTGACTACTCATTCATACATGGAAGAGCGCAGAGGCTGTCTGT 4709
QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520
Db GACAGAGGGCCCATCATGTTCTGGCGCTTAAGTGTATTTGGAACAGTACACATCTA 4769
QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluLysLeuLysLeuLysAspTrpIleGluAsn 1540
Db AAGAGTAAGTTTGTGTAAGGCTACTTTTGGAAATTAATAAAGCACTGGATAGAAAC 4829
QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560
Db CTGGAATAAGTAGTCGCTTCAAGAGAAATCAATATATTTTCCCAATATGCAAGCCGCA 4889
QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580
Db GAGAGCTTTTCTTCTATTTTGGCTTTTAAATTCAAAAGAAAGATGTCAGTCTCTCTCA 4949
QY 1581 GlnSerPhePheLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600
Db CAATCCTTCGCTAAGTTGGAAGAACAAACGCACTTTTGGCATTGAAGAATACAGCTTT 5009
QY 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620
Db TCTCAAGCAACCTCGAACAGGTTTTTGTAGAACTTACTAAAGAGCAAGAGAGAGAGAT 5069
QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640
Db AACAGCTGTGGAGCTTTGGCCAGCACTCTCTGTGGAAAGAGCAGCAGAGGATAGAGTA 5129
QY 1641 ValPhe 1642
Db GTATTT 5135

```

## RESULT 15

```

HSM806823      4078 bp      mRNA      linear      PRI 28-AUG-2003
LOCUS      Homo sapiens mRNA; cDNA DKFP686K07118 (from clone DKFP686K07118).
DEFINITION      BX640745
ACCESSION      BX640745.1  GI:34364843
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4078)
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
CONSRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing Consortium of the
German Genome Project.
This clone (DKFZp686K07118) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
source
1..4078
location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /map="17q24.3"
    /clones="DKFZp686K07118"
    /tissue_type="human fetal kidney"
    /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
    DH10B; sites SfiIA + SfiIB"
    /dev_stage="fetal"
    1..4078
    /gene="DKFZp686K07118"
    <1..1296
    /gene="DKFZp686K07118"
    /note="ATP-binding cassette A5, N-terminus truncated,
    differentially spliced"
    /codon_start=1
    /product="hypothetical protein"
    /protein_id="CAE45856.1"
    /db_xref="GI:34364844"
    /translations="VFVFVTVAFGFIGLMILTESPKSLVWLFSPFCHCTFVIGIAQ
    VMLEDNEGASPSNLTAGPYPLIIRIMLTANSIPVLLAVLYDQVIRPGEGLRESS
    LYFLKSYWSKSRNVEESEGNGNISFSEIIEVSEFVGKEAIRISGIQKYRK
    KGENVEARNLSFDIEGQITALLHSGTGKSTLMNLGCLPPSPDGPSFIYGHVSE
    IDEMFEARKMIGICPDQIDHFVLTVEENLSILAKGIPANNIOEVOKVLLDDMQ
    TIKDQAKLSGGQKRLSLGIAVLGNPKILLDEPTAGMDPCSRHIVNLLKYRAN
    RVTVFSHFWDDEAILADRAKAVISQGMKCVGSMFLKWKWGIYRLSVIDKYCATE
    SLASLVKQHTPGATLLQQNDQQLVYSILPFKMDKFSGLA"
    4041..4046
    /polyA_signal
    /gene="DKFZp686K07118"
    4056
    /polyA_site
    /gene="DKFZp686K07118"
    ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4078
Score: 6477.00 Matches: 1272
Percent Similarity: 95.64% Conservatives: 0
Best Local Similarity: 95.64% Mismatches: 1
Query Match: 76.87% Indels: 57
DB: 9 Gaps: 3
US-10-090-458-5 (1-1642) x HSM806823 (1-4078)
Qy 330 ValGluPhePheValThrValAlaPheGlyPheIleGlyLeuMetIleIleGlu 349
Db 1 GTTGAATTTTGTACTGGCTTTTGGATTTATTGGCCTTATGATAATCCTCATAGAA 60
Qy 350 SerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCysThrPheValIle 369
Db 61 AGTTTCCCAATCGTTAGTGTGGCTTTTCAGTCCCTTCTGTCACGTACTTTGTGATT 120
Qy 370 GlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSerPheSerAsnLeu 389

```

```

121 GGTATTGCACAGGTCATGCTTTTGAAGATTTTAATGAAGTGTCTTCAATTTCAAATTTG 180
Qy 390 ThrAlaGlyProTyrProLeuIleIleThrIleIleMetLeuThrLeuAsnSerIlePhe 409
181 ACTGCAGGCCCATATCCTCTTAATTATTAACAATTATCATGCTCACACTTAATAGTATATC 240
Qy 410 TyrValLeuLeuAlaValTyrLeuAspGlnValIleProGlyGluPheGlyLeuArgArg 429
241 TATGTCCTCTTTGGTGTCTATCTTCAAGTCATTCAGGGGAATTTGGCTTACGGAGA 300
Qy 430 SerSerLeuTyrPheLeuLysProSerTyrTrpSerLysSerLysArgAsnTyrGluGlu 449
301 TCATCTTTATATTTCTGAAGCCTTCATATTGGTCAAGAGCAAAAGAAATATGAGGAG 360
Qy 450 LeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGluIleIleGluProValSer 469
361 TTATCAGAGGCAATGTTTAATGGAATATTAGTTTAAATTTAGTGAATATTAGGCCAGTTCT 420
Qy 470 SerGluPheValGlyLysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLys 489
421 TCAGATTTTGTAGGAAGAAGCCTAAGAATTAGTGTATTTCAGAGACATACAGAAAG 480
Qy 490 LysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIle 509
481 AAGGCTGAAATGTGGAGGCTTTGAGAAATTTGTCAATTGACATATATGAGGGTCAGATT 540
Qy 510 ThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGly 529
541 ACTGCCCTTACTTGGCCACAGTGGCAACAGGAAGAGTACATTGATGATATTTCTTTTGGGA 600
Qy 530 LeuCysProProSerAspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAsp 549
601 CTCGCCCCACCTCTCTGATGGGTGTTGCATCTATATATGACACACAGAGTCTCAGAAATAGAT 660
Qy 550 GluMetPheGluAlaArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAsp 569
661 GAAATGTTTGAAGCAGAGAAATAATGATGGCAATTTGTCCACAGTTAGATATACACTTTGAT 720
Qy 570 ValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsn 589
721 GTTTTGACAGTAGAAGAAATTTATCAATTTTGGCTTCAATCAAAAGGATACCAGCCAAC 780
Qy 590 AsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLysAsp 609
781 AATATAATACAGAAGTGCAGAGGTTTTACTAGATTTAGACATGCACACTATCAAGAT 840
Qy 610 AsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaVal 629
841 AACCAAGCTTAATAAATTAAAGTGGTGGTCAAAAAGAAAGAGCTGTCTATTAGGAATTGCTGT 900
Qy 630 LeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSer 649
901 CTTGGGAACCCAAAGATACCTGCTAGATGAACCAACAGCTGGAATGGACCCCTGTTCT 960
Qy 650 ArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSer 669
961 CGACATATTGTATGGATCTTTTAAATACAGAAAGCCATCGGTGCACAGTGTTCAGT 1020
Qy 670 ThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGly 689
1021 ACTCATTTTCATGGATGAAGCTGACATTTCTTCAGATAGGAAGCTGTGATATCAAGGA 1080
Qy 690 MetLeuLysCysValGlySerSerMetPheLeuLysSerLysTyrTpGlyIleGlyTyrArg 709
1081 ATGCTGAATGTGTTGGTTCTTCAATGTTCTCTCAAAAGTAATGGGGATCGGTACCGC 1140
Qy 710 LeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGln 729
1141 CTGAGCATGTACATAGACAAATATTGCCACAGAAATCTCTTTCTTCACTGGTAAACAA 1200
730 HisIleProGlyAlaThrLeuLeuGlnAsnAspGlnGlnLeuValTyrSerLeuPro 749

```

Db 1201 CATATACCTGGAGCTACTTTATTACACAGAAATGACCAACAACTTGTGTATAGCTTGGCT 1260  
 Qy 750 PheLysAspMetAspLysPheSerGly----- 758  
 Db 1261 TTCAAGACATGGACAAATTTTCAGGAATGCTGTATAGACAAAAGAGTGTATTATGATGT 1320  
 Qy 759 LeuPheSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMet 778  
 Db 1321 TTGTTTTCTGCCCTAGACAGTCATTCAAATTTGGGTGTCATTCTTATGGTGTTCATG 1380  
 Qy 779 ThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGluLeuAspGlnAlaAsp 798  
 Db 1381 ACGACTTTGGAAGACGTATTTTAAAGCTAGAAGTTGAACGAAATTTGACCAAGCAGAT 1440  
 Qy 799 TyrSerValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGlu 818  
 Db 1441 TATAGTGTATTACTACAGAGCCACTGGAGAGAAATGATTCAAAATCTTTTGATGA 1500  
 Qy 819 MetGlnGlnSerLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSer 838  
 Db 1501 ATGGAACAGAGCTTACTATTCTTTCTGAAACCAAGGCTTCTTAGTGACACCATGAGC 1560  
 Qy 839 LeuTrpLysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGlu 858  
 Db 1561 CTTTGGAAACACAGATGTATACAAATAGCAAGTTTCATTCTTACCTTGAACCTGAA 1620  
 Qy 859 SerLysSerValArgSerValLeuLeuLeuLeuLeuPhePheThrValGlnIlePhe 878  
 Db 1621 AGTAAATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTT 1680  
 Qy 879 MetPheLeuValHisSerPheLysAsnAlaValValProIleLysLeuValProAsp 898  
 Db 1681 ATGTTTTGGTTCATCACTCTTTTAAAAATGCTGTGTTCCCATCAACTGTTCAGAC 1740  
 Qy 899 LeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGln 918  
 Db 1741 TTATATTCTTAAACCTGGAGACAAACCATAAATACAAACCAAGTCTGCTTCTTCAA 1800  
 Qy 919 AsnSerAla-----AspSerAspIleSerAspIleSerPhePhe 933  
 Db 1801 AATTCTGCTGTGAGAGTGTGAAGACTCAGATATCAGTGATCTTATTAGCTTTTTCAC 1860  
 Qy 933 rSerGlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHi 953  
 Db 1861 AAGCCAGACATATGTTGAGATGATTAATGACAGTACTATGATCCGTGGCTCCCA 1920  
 Qy 953 sSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaValPheAs 973  
 Db 1921 TAGTGGCGCTTAAATGTGATGATTCAGAAAGGACTATGTTTTTGCAGCTGTTTCAA 1980  
 Qy 973 nSerThrMetValTyrSerLeuProIleLeuValAsnIleLeuSerAsnTyrTyrLeuTy 993  
 Db 1981 CAGTACTATGTTTATCTTCTTACTATATAGTGAATATCATTAGTAACTACTATCTTTA 2040  
 Qy 993 rHisLeuAsnValThrGluThrIleGlnIleTrpSerThrProPhePheGlnGluLeu 1013  
 Db 2041 TCAITTTAAATGTACTGAACCATCCAGATCTGGAGTACCCCATCTTCTTCAAGAAATTC 2100  
 Qy 1013 rAspIleValPheLysIleGluLeuTyrPheGlnAlaLeuLeuGlyIleIleValTh 1033  
 Db 2101 TCAATAGTTTTTAAATGTAGCTGTATTTTCAAGAGCTTTGCTTGGAAATCATGTGTAC 2160  
 Qy 1033 rAlaMetProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrTh 1053  
 Db 2161 TGCAATGCCACTTACTTTTGCATGGAAATGCAGAGAAATCATTAAGATCAAGCTTATAC 2220  
 Qy 1053 rGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValas 1073  
 Db 2221 TCAACTTAAACTTTCAGGTCTTTTGGCCATCTGCATATGGATTTGGCAAGCTGTGTGTA 2280  
 Qy 1073 pIleProLeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTy 1093  
 Db 2281 TATCCCTTATTTTTTATCATCTTATTTTGTATGCTAGGAAGCTTACTGGCATTTTCA 2340

Qy 1093 rGlyLeuTyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrVa 1113  
 Db 2341 TGGATTATATTTTATACGTAAAGTTCCTTGCTGTGTTTTTTCCTTATTTGCTTATGT 2400  
 Qy 1113 lProSerValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnTh 1133  
 Db 2401 TCCATCAGTTATCTGTTCACTATATGCTTCTTCACCTTTAAGAAAAATTTTAAATAC 2460  
 Qy 1133 rLysGluPheTrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGl 1153  
 Db 2461 CAAGAATTTTGGTCATTTATCTATTCTGTGGAGCGTGTGGCTGTGATTTCGATCATCTGA 2520  
 Qy 1153 uIleThrPhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIle 1173  
 Db 2521 AATAACTTCTTTATGGGATACAAATGCAATTCATCTTCATTTATGCTTTTGTATCAT 2580  
 Qy 1173 eIleProIleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAs 1193  
 Db 2581 CATTCCAATCTATCATTCTAGTGTGCTGATTTCTTTCTATAAAGATTCTTCTGGAAGA 2640  
 Qy 1193 nValArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSe 1213  
 Db 2641 TGTACGAAAAAATGGAACACTTAAATCCATGGATAGGCTTTCAGTAGCTGTATATATC 2700  
 Qy 1213 rProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGl 1233  
 Db 2701 GCCTTACCTGAGTGTGACTGTGGATTTCTCTTCAATACTATGAGAAAAATATGG 2760  
 Qy 1233 yGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnAr 1253  
 Db 2761 AGGCAGATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTAAATAATAG 2820  
 Qy 1253 gLysLeuProGluProProAspAsnGluAspGluAspValLysAlaGluArgLe 1273  
 Db 2821 GAAGCTTCCAGAACCCAGACCAATGAGGATGAAGATGAAGATGTCAAAGCTGAAGACT 2880  
 Qy 1273 uLysValLysGluLeuMetGlyCysGlnCysCysGluGluLysProSerIleMetValSe 1293  
 Db 2881 AAAGTCTAAAGAGCTGATGGGTGCCAGTGTGTGGAGAGAAACCATCCATTATGTCAG 2940  
 Qy 1293 rAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSerArgLysValLysLy 1313  
 Db 2941 CAATTTGCATAAAGAAATATGATGACAGAAAGATTTTCTCTTTCAAGAAAAGTAAAGAA 3000  
 Qy 1313 sValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIleLeuGlyLeuLeuGl 1333  
 Db 3001 AGTGGCAACTAAATACATCTCTTTCTGTGTAAGAAAGAGAGATCTTAGGACTATTGGG 3060  
 Qy 1333 yProAsnGlyAlaGlyLysSerThrIleIleAsnIleLeuValGlyAspIleGluProTh 1353  
 Db 3061 TCCAAATGGTGTGCGCAAAAGCACAATTAATTAATATCTGTTGGTGTATATTGAACCAAC 3120  
 Qy 1353 rSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAspSerLe 1373  
 Db 3121 TTCAGCCAGGTATTTTAGGAGATTTCTTCAGACCAAGTGAAGATGATGATTCAT 3180  
 Qy 1373 uLysCysMetGlyTyrCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGl 1393  
 Db 3181 GAAGTGTATGGTTACTGCTCTCAGATAAACCCCTTTGCGCCAGATATACATTGCGAGA 3240  
 Qy 1393 uHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIl 1413  
 Db 3241 ACATTTGAAATTTATGAGCTGTCAAAGGAATAGTGCAGTGCATGAAGAAAGTCTCAT 3300  
 Qy 1413 eSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLe 1433  
 Db 3301 AAGTCGAATAACACATGCACTTGAITTTAAAGAAACATCTTCAGAAGACTGTAAAGAAACT 3360  
 Qy 1433 uProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIl 1453  
 Db 3361 ACCTGAGGAATCAAAACGAAAGTTGTGTTTGTCTTAAGATGCTAGGAAATCTTCAGAT 3420

Search completed: April 13, 2004, 00:30:47  
Job time : 9921 secs



GenCore version 5.1.6  
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 12, 2004, 17:58:28 ; Search time 902 Seconds  
(without alignments)  
7733.416 Million cell updates/sec

Title: US-10-090-458-5

Perfect score: 8426

Sequence: 1 MSTAIREVGVWQRTLLK.....CGTLNSTLWERTQEDRVVP 1642

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10090458/runat\_08042004\_121552\_25446/app\_query.fasta\_1.1799  
-DB=N\_Geneseq\_29Jan04 QPMT=fastap SUFFIX=ring -MINMATCH=0.1 -LOOFCU=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAIRIX=BLOSUM62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10090458 @CEN 1 1 636 @runat\_08042004\_121552\_25446 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8426	100.0	5475	7 ABS57751	Ab57751 cDNA enco
2	8422	100.0	6369	6 AAD33648	Aad33648 Human TRI
3	8422	100.0	6525	6 ABN89594	Abn89594 Human ATP
4	8395	99.6	4929	6 AAD37618	Aad37618 Human tra
5	8395	99.6	5262	6 AAD37620	Aad37620 Human tra
6	8309	98.6	4917	7 ABS57750	Ab57750 Coding se
7	8309	98.6	5463	7 ABS57749	Ab57749 cDNA enco
8	8112	96.3	4785	6 AAD37619	Aad37619 Human tra

9	4847.5	57.5	3268	7	ADA53416	Ada53416 Human cod
10	4282	50.8	2723	9	ADC51606	Adc51606 Human mac
11	3259	38.7	5797	8	ADA20300	Ada20300 Human ATP
12	3234	38.4	6181	6	ABN89597	Abn89597 Human ATP
13	3229	38.3	5722	7	ABZ23292	Abz23292 Human ABC
14	3200.5	38.0	4875	7	ABS57519	Ab57519 Human ABC
15	3200.5	38.0	5018	7	ABS57518	Ab57518 Human ABC
16	3197.5	37.9	5981	6	ABN89596	Abn89596 Human ATP
17	3195	37.9	5311	9	AAD29639	Aad29639 Mouse tum
18	3192	37.9	5680	7	ABZ23292	Abz23292 Human ABC
19	3124	37.1	5211	7	AAD47363	Aad47363 Human tra
20	3124	37.1	5296	6	ABN89595	Abn89595 Human ATP
21	3117	37.0	5332	6	ABQ77735	Abq77735 Human ABC
22	3108	36.9	4743	8	ABQ77571	Abq77571 Drug tran
23	3108	36.9	5677	6	ABT10141	Abt10141 Human bre
24	3107.5	36.9	3928	6	ABK35707	Abk35707 cDNA sequ
25	3065	36.4	5149	6	AAD33675	Aad33675 Human TRI
26	3044	36.1	5622	6	AAD33657	Aad33657 Human TRI
27	2991.5	35.5	5065	6	AAD33650	Aad33650 Human TRI
28	2982	35.4	4798	6	ABV99425	Abv99425 Human NOV
29	2971	35.3	5846	7	AAD49503	Aad49503 Human TRI
30	2961.5	35.1	4766	6	ABV99426	Abv99426 Human NOV
31	2824.5	33.5	4727	6	AAL44690	Aal44690 Human tra
32	2660	31.6	1548	6	ABK35706	Abk35706 cDNA sequ
33	2333	27.7	2481	7	ABZ35938	Abz35938 Human sec
34	2229.5	26.5	6848	7	ABX34749	Abx34749 Human mdd
35	1989	23.5	3806	7	ABX74407	Abx74407 Human CDN
36	1760.5	20.9	4345	5	AAS66371	Aas66371 DNA encod
37	1677	19.9	1346	6	ABL64459	Ab164459 Stomach c
38	1677	19.9	1346	6	ABL63763	Ab163763 Breast ca
39	1677	19.9	1346	6	ABN96909	Abn96909 Gene #340
40	1674	19.9	2645	7	ADA53722	Ada53722 Human cod
41	1637	19.4	5971	5	AAS70348	Aas70348 DNA encod
42	1555.5	18.5	6470	4	AAS74445	Aah57445 Human lun
43	1553.5	18.4	6491	3	AZ94761	Az94761 Human ATP
44	1533.5	18.4	6451	6	ABL66808	Ab166808 Lung canc
45	1533.5	18.4	6491	6	ABK84439	Abk84439 Human CDN

## ALIGNMENTS

### RESULT 1

ID ABS57751 standard; cDNA; 5475 BP.

XX AC ABS57751;

XX DT 04-FEB-2003 (first entry)

XX DE cDNA encoding novel human ATP binding cassette ABCA5 transporter #2.

XX KW Human; ATP binding cassette; ABC; ABCA5; transporter;

XX KW neurotoxin transport; beta-amyloid peptide; chromosome mapping;

XX KW blood brain barrier transport; tissue typing; predictive medicine;

XX KW ABCA5 mediated disorder; ABCA5 related disorder; gene therapy; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 114..5042

XX FT /\*tag= a

XX FT /product= "ABCA5"

XX FT /note= "ATP binding cassette (ABC) A5 transporter"

XX FN US2002123107-A1.

XX PD 05-SEP-2002.

XX PF 01-MAR-2002; 2002US-00090458.

XX XX 02-MAR-2001; 2001US-0272885P.

XX XX (ACTI-) ACTIVE PASS PHARM INC.

XX Chen H, Kilinski L, Le Bihan S;  
 XX WPI; 2003-066798/06.  
 DR P-PSDB; ABG72424.  
 XX Novel isolated ATP binding cassette transporter family polypeptide,  
 PT ABCA5, useful for treating disorders associated with aberrant or unwanted  
 PT ABCA5 transporter expression or activity.  
 XX Claim 2; Page 39-42; 52pp; English.  
 XX The invention describes an isolated ATP binding cassette (ABC)  
 CC transporter family polypeptide (I), designated ABCA5. (I) or the  
 CC polynucleotide encoding it (II) are useful as targets for developing  
 CC modulating agents to regulate a variety of cellular processes,  
 CC particularly the transport of neurotoxic molecules, e.g., beta-amyloid  
 CC peptide (Aβ), across cell membranes or, e.g., the blood brain barrier  
 CC (BBB), as targets for developing modulating agents of multi-drug  
 CC resistance, as diagnostic and therapeutic tools, or to treat disorders  
 CC associated with aberrant or unwanted ABCA5 transporter expression or  
 CC activity. (I), (II) or a host cell (III) expressing (II) are useful in  
 CC screening assays, detection assays (e.g., chromosomal mapping, tissue  
 CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,  
 CC prognostic assays, monitoring clinical trials and pharmacogenomics), and  
 CC in methods of treatment (e.g., therapeutic and prophylactic). (I) or  
 CC (III) are useful as reagents or targets in assays applicable to treatment  
 CC and diagnosis of ABCA5-mediated or related disorders. (I) is useful to  
 CC screen for naturally occurring ABCA5 substrates; to screen for drugs or  
 CC compounds which modulate ABCA5 activity; as a bait protein in a yeast two  
 CC -hybrid or three-hybrid assay; and to identify other proteins which bind  
 CC to or interact with ABCA5. (II) is useful in: gene therapy; to detect  
 CC ABCA5 mRNA or a genetic alteration in an ABCA5 gene; to modulate ABCA5  
 CC activity; to locate gene regions associated with genetic disease or to  
 CC associate ABCA5 with the disease, to identify an individual from a minute  
 CC biological sample (tissue typing), and to aid in forensic identification  
 CC of the biological sample. This sequence encodes a novel human ATP binding  
 CC cassette (ABC) A5 transporter.  
 XX Sequence 5475 BP; 1705 A; 917 C; 1048 G; 1804 T; 0 U; 1 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 0 Length: 5475  
 Score: 8426.00 Matches: 1642  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 7 Gaps: 0

US-10-090-458-5 (1-1642) x ABS57751 (1-5475)

QY 1 MetSerThrAlaIleArgGluValTyrArgGlnThrArgThrLeuLeuLeuLys 20  
 DB 114 ATGTCACGTCAATTAGGAGGTAGAGTTGGACAGACCCAGAACACTTCTACTGAAG 173  
 QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGluLeuPheProLeu 40  
 DB 174 AATTACTTAATTAATGCAGAACCAAAAGAGTAGTGTTCAGAGAAATCTTTTCCACTA 233  
 QY 41 PhePheLeuPheTrpLeuLeuLeuLeuSerMetMetHisProAsnLysLysTyrGluGlu 60  
 DB 234 TTTTITTTTATTTTGGTAAATTAATTAATTCAGTCATGATCCCAATTAAGAAATATGAAGA 293  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
 DB 294 GTGCCTAATATAGAACTCAATCTATGGACAGTTTACTCTTCTAATCTAATTTCTTGA 353  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 DB 354 TATACTCCAGTCACTAATATTACAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA 413  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120

DB 414 CCTGATGTCAATTAATTACTGAAGAATATACAAATGAAAAAAGAAATGTTAAACATCCAGTCTC 473  
 QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
 DB 474 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAAGACTCCCATGTCTTATGAACCTCGT 533  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 DB 534 TTTTITTTCTGATATGATTCAGTATCTTCTAATTAATGATGATTCAGAGCTGGCTGTCA 593  
 QY 161 LysSerCysGluAlaIleGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 DB 594 AATATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTACAGTTTTCAGAGCATCCATA 553  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGluSerThr 200  
 DB 654 GATGCTGCCATATACAGTTGAAGCAATGTTTCTCTTTGGAGGAGCTGGAGTCAACT 713  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
 DB 714 AAAGCTGTTTATTTATGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGTA 773  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 DB 774 ATTTTATATACCTAGTTATAGCATTTTCACTTTTGGTACTTTTGGCAATTCATATC 833  
 QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 DB 834 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATATATGGGACTTCATGATACT 893  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 DB 894 GCCTTTTGGCTTCT 953  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
 DB 954 ATGGCAGTCAATTCGACAGCTTCTTTGTTATTTTCTCAAGTAGCAGCAATTTGATATTT 1013  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 DB 1014 CTGCTTTTCTCTTTATGATATATCATCTGATATTTTCTGTTTATGCTGACACCTCTT 1073  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 DB 1074 TTTAAAAAATCAAAACATATGGGAATAGTTGAATTTTGTACTGTGGCTTTTGGATTT 1133  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 DB 1134 ATGGCCCTTATGATAATCTCTCATAGAAAGTTTCCCAATTCGTTAGTGGCTTTTCAGT 1193  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 DB 1194 CCITTCGTGTCACGTACTTTTGTGATGTTGTTATGTCAGGTCATGATTTAGAAGATTTT 1253  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 DB 1254 AATGAAGTCTCTCATTTTCANATTTGACTGCGAGGCCATATCCTCTAATTTATTAACAT 1313  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 DB 1314 ATCATGTCTACACTTAATAGTATATCTATGCTCTCTGCTGCTATCTTGTATCAAGTC 1373  
 QY 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 DB 1374 ATTCAGGGGAATTTGGCTTACGAGATCATCTTATATTTTCTGAGGCCCTTCATATTGG 1433  
 QY 441 SerLysSerLysArgAsnTyrGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 DB 1434 TCAAAGACAAAGAAATTAATGAGGATATCAGAGGCAATGTTAATGGAATATTAGT 1493  
 QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
 DB 1494 TTTAGTGAAATTTATGAGCCAGTTTCTTCAAGATTTTGTAGAAAAAGAACCAAGAATTT 1553

QY 481 SerGlyLeuGlnLysThrTyrArgLysLysGlyGluAenValGluAlaLeuAraAsnLeu 500  
Db 1554 AGTGGTATTTCAGAGACATACAGAAAGAGGGTGAATAATGTGGAGGCTTTGAGAAATTTG 1613  
QY 501 SerPheAspIleTyrGluGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
Db 1614 TCATTGGACATATAGAGGGTCAGAAATTAAGTTCCTTACCTTGGCCACAGTGGAAAG 1673  
QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
Db 1674 AGTACATTGATGAATATTCTTTGTGGACTCTGCCACCTCTCTGATGGGTTTGCATCTATA 1733  
QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db 1734 TATGACACAGAGTCTCGAATAGATGAATGTTTGAAGCAGAGAAATGATTGGCAATT 1793  
QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAenLeuSerIleLeu 580  
Db 1794 TGTCACACAGTTAGATATACACTTTGATGTTTGACAGTAGAAGAAATTTATCAATTTG 1853  
QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleLeuGlnGluValGlnLysValLeuLeu 600  
Db 1854 GCTTCAATCAAGGATACCAAGCAACAATATATACAAAGAGTGCAGAGGTTTACTA 1913  
QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
Db 1914 GATTTAGACATGCACACTATCAAGATTAACAGCTAATAAATTAAGTGGTGCACAAA 1973  
QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAenProLysIleLeuLeuLeuAspGlu 640  
Db 1974 AGAAGAGTGTCAATTAGGAATTTGCTTCTGGGAACCCAAAGATACTGCTAGATGAA 2033  
QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAenLeuLeuLysTyrArg 660  
Db 2034 CCACAGCTGGAAATGGACCCCTGTTCTCGACATATTGTATGTAATCTTTTAAATACAGA 2093  
QY 661 LysAlaAenArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 2094 AAAGCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATCTTGCA 2153  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2154 GATAGGAAGCTGTATATATCAAGAGAAATGCTGAATGTGTGTGTTCTTCAATGTTCCCTC 2213  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2214 AAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACA 2273  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db 2274 GAATCTCTTCTTCACCTGGTTAAACAAACATATACCTCGAGCTACTTTATTACACAGAA 2333  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 2334 GACCAACAACTTGTTGATAGCTTGCTTTCAGGACATGGCAAAATTTTCAGGTTGTTT 2393  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2394 TCTGCCCTAGACAGTCATTTCAAAATTTGGGTGTCATTTCTTATGGTGTTCATGACGACT 2453  
QY 781 LeuGluAspValPheLeuLysLeuValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2454 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATAGT 2513  
QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2514 GTATTTACTCAGCAGCCACTGGAGAGAAATGATTCAAAATCTTTGATGAATGGAA 2573  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 2574 CAGAGCTTACTTATTCTTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGG 2633

QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrIleuLysArgGluSerLys 860  
Db 2634 AAACAACAGATGATATCAATAGCAAAAGTTTCATTTCTTTACCTTGAAACGTGAAAGTAA 2693  
QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAGTTTCAGATTTTTATGTTT 2753  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 2754 TTGGTTTCATCTCTTTTAAATATGCTGGTCCCATCAAACTGTTCCAGACTTATAT 2813  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
Db 2814 TTTCTAAACCTGGAGACAAACACATAAATCAAAACCAAGTCTGCTCTTCAAAATCT 2873  
QY 921 AlaAspSerAspIleSerAspLeuLeuSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 2874 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAGCCAGAACATAATGGTGACG 2933  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 2934 ATGATTAATCAGAGTACTATGTCGGTCCCATAGTGGGCTTTTAAATGTGATG 2993  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 2994 CATTCAAGAAAGGACTATGTTTTGCAGAGCTGTTTCAACAGTACTATGTTTATTCTTTA 3053  
QY 981 ProIleLeuValAsnIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 3054 CCTATATTAGTGAATATCATTAAGTACTATCTTTTATCATTTAAATGTGTAACC 3113  
QY 1001 IleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 3114 ATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTAAGTATAGTATTTTAAATTTAG 3173  
QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 3174 CTGTATTTTCAAGCAGCTTTGCTTGGATCAATGTGTACTGCAATGCCCTTATTTTATCAT 3233  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3234 ATGGAATATGCAAGAAATCATTAAGTCAAGCTTTATCTCAACTTAAACTTTTCAGGCTTT 3293  
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAspIleProLeuPhePheIle 1080  
Db 3294 TTGCCATCTGCATATTGGATTGGACAAGCTGTGTGTATATCCCTTATTTTATCAT 3353  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3354 CTATTTTGTATGCTAGGAAGCTTATGGCAATTTCAATATGATGATATATTTTATCTGTA 3413  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 3414 AAGTCTCTGCTGTGTTTTTGGCTTATGTTGTTATGTTTCCATCAGTTATTTCTGTTCACT 3473  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 3474 TATATTGCTCTCTTCCCTTTAAGAAATTTTAAATATACAAAGAAATTTTGGTCATTATTC 3533  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3534 TATCTGTGGCAGCGTTGGCTTGTATTCGAATCACTCAATCAATTTCTTTTATGGATAC 3593  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 3594 ACAATTTGCAACTATTCTTCATTATGCTTTTGTATCATCATTTCCCAATCTATCCACTTTTA 3653  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
Db 3654 GGTGCTGCTGATTTCTTTCATAAAGATTTCTTGGAGAGATGTACGAAAAATTTGGACACC 3713  
QY 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220

```

Db 3714 TATAATCCATGGATAGGCTTTCAGTAGCTGTATATCGCTTACCTGCAGTGACTG 3773
Qy 1221 TrpIlePheLeuLeuGlnTyrTyrGluYsIysTyrGlyArgSerIleAeGlyLeuAsp 1240
Db 3774 TGGATTTTCCCTCTTACAAATATATGAGAAAATAATGGAGCAGATCAATAAGAAAAGAT 3833
Qy 1241 ProPhePheArgAsnLeuSerThrIysSerIysAsnArgLysLeuProGluProProAsp 1260
Db 3834 CCCTTTTTCAGAAACCTTTCAACGAAGTCTAAANAATAGGAAGCTTCCAGAACCCACAGAC 3893
Qy 1261 AsnGluAspGluAspGluAspVallysalagluargleuYsVallyGluLeuMetGly 1280
Db 3894 AATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGGT 3953
Qy 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300
Db 3954 TGCCAGTGTGTGAGGAGAACCCATCCATTATGTCAGCANATTTGCATAAAGATATGAT 4013
Qy 1301 AspIysLysAspPheLeuLeuSerArgLysVallysalalathrLysTyrIleSer 1320
Db 4014 GACAAAGAAAGATTTTCTCTTTCAAGAAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 4073
Qy 1321 PheCysValLysLysGlyGluLysLeuGluLeuGlyProAsnGluValaglyLysSer 1340
Db 4074 TTCTGTGTGAAAAGAGAGAGATCTTAGACATATTTGGTCCCAATGGTGTGGCAAAAGC 4133
Qy 1341 ThrIleLeuAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360
Db 4134 ACAATTAATTAATTTCTGTTGGTGGATATTGAACCAACTTCAGGCCAGGTATTTTAGGA 4193
Qy 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380
Db 4194 GATTATTTCTTCAGACAAAGTGAAGATGATGATTCAGTGAAGTGTATGGTTACTGCTCT 4253
Qy 1381 GlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400
Db 4254 CAGATAAACCCCTTTGTGGCCAGATACATACATTGCAGGAACAATTTTGAATTTATGGAGCT 4313
Qy 1401 VallyGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420
Db 4314 GTCAAGAGGAATGAGTGAAGTGAATGAAGAGATCATAGTGAATGAATGAATGAATGACAT 4373
Qy 1421 AspLeuLysGluHisLeuGlnLysThrVallyLysLeuProAlaGlyIleLysArgLys 1440
Db 4374 GATTTAAAGAAACATCTTCAGAAAGACTGTAAAGAACTACCTGCAGGAATCAAAAGAAAG 4433
Qy 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460
Db 4434 TTGTGTTTTGCTCTAAGTATGCTAGGGAATCTCAGATTACTTTGCTAGATGAACCATCT 4493
Qy 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLys 1480
Db 4494 ACAGGTATGATCCCAAGCCCAACAGACACATGTGGCGAGCAATTCGAATTCGATTTAAA 4553
Qy 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500
Db 4554 AACAGAAAGCGGCTGCTATTCGACCACTCACTATATGAGAGGAGGAGGAGGCTGTCTGT 4613
Qy 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520
Db 4614 GATCGATGATCTATCATGTGTCTGGCAGTTAAGATGTATCGGNACAGTACACACTA 4673
Qy 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTyrIleGluAsn 1540
Db 4674 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTTAAATTTGAAGAGCTGATGATGAAAC 4733
Qy 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560
Db 4734 CTAGAGTAGACCGCTCTCAAAGAGAAATTCAGTATATTTTCCCAATTCGAGCGCTCAG 4793
Qy 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValSerLeuSer 1580

```

```

Db 4794 GAAAGTTTTTCTTCTATTGGCTTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTCA 4853
Qy 1581 GlnSerPhePheLeuLeuGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600
Db 4854 CAATCTTTTTTAACTGGAGAGACTAAACATGTTTTTGCATTTGAAGATATAGCTTT 4913
Qy 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620
Db 4914 TCTCAAGCAACATTCGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGGAGAGAT 4973
Qy 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyrTyrGluArgThrGlnGluAspArgVal 1640
Db 4974 ATAGTGTGGAACTTTTAAACAGACACATTTGTGGTGGACGACACACAAAGATAGATGA 5033
Qy 1641 ValPhe 1642
Db 5034 GTATTT 5039

RESULT 2
AAD33648 standard; cDNA; 6369 BP.
ID AAD33648;
AC AAD33648;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human TRICH-3 cDNA.
XX
KW Human; transporter and ion channel; TRICH-3; transport disorder; angina;
KW amyotrophic lateral sclerosis; cystic fibrosis; diabetes; neuromuscular disorder;
KW cardiac disorder; polymyositis; diabetes; neurological disorder; cancer;
KW depression; schizophrenia; anaemia; Wilson's disease; Cushing's disease;
KW cell proliferated disorder; infertility; arteriosclerosis; gene therapy;
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; allergy;
KW myasthenia gravis; multiple sclerosis; metabolic disorder; hypertension;
KW acquired immune deficiency syndrome; immunological disorder; scleroderma;
KW endocrine disorder; autoimmune thyroiditis; rheumatoid arthritis; goitre;
KW cardiac myopathy; amnesia; toxic myopathy; Addison's disease; infection;
KW epilepsy; mental disorder; myocarditis; Crohn's disease; Grave's disease;
KW muscle disorder; stroke; dementia; anxiety; AIDS; asthma; cirrhosis;
KW gene, ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
CDS 1238..6166
    /tag= a
    /product= "Human TRICH-3 protein"
XX
W0200212340-A2.
XX
14-FEB-2002.
XX
01-AUG-2001; 2001WO-US024217.
XX
03-AUG-2000; 2000US-0223269P.
XX
10-AUG-2000; 2000US-0224456P.
XX
18-AUG-2000; 2000US-0226410P.
XX
25-AUG-2000; 2000US-0228140P.
XX
31-AUG-2000; 2000US-0230067P.
XX
08-SEP-2000; 2000US-0231434P.
XX
(INCY-) INCYTE GENOMICS INC.
XX
Yue H, Thornton M, Ramkumar J, Tang YT, Azimzai Y, Baughn ME;
PI Yang J, Yao MG, Lal P, Walia NK, Gandhi AR, Hafalia AJA, Nguyen DB;
PI Patterson C, Elliott VS, Tribouley CM, Lu DAM, Xu Y, Reddy R;
PI Hernandez R, Borowsky ML, Lo TP, Lu Y, Policky JL, Greene BD;
PI Sanjanwala MS, Raumann BE, Burford N, Ison CH, Lee EA, Ding L;
PI Das D, Kallick DA, Khan FA, Seilhamer JJ;
XX
WPI; 2002-206330/26.
DR P-PSDB; AAE21159.

```

XX New human transporters and ion channels polypeptides and polynucleotides  
 PT for diagnosing, preventing or treating transport, neurological, muscle,  
 PT immunological and cell proliferative disorders.  
 XX

PS Claim 77; Page 203-205; 230pp; English.

XX The invention relates to human transporter and ion channel polypeptides  
 CC designated TRICH and nucleic acid molecules encoding such polypeptides.  
 CC TRICH sequences are useful for diagnosis, treatment and prevention of  
 CC transport, muscle, neurological, immunological and cell proliferative  
 CC disorders. Transport disorders include akinesia, amyotrophic lateral  
 CC sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular  
 CC dystrophy, diabetes mellitus, diabetes insipidus, myasthenia gravis,  
 CC myocardiitis, prostate cancer, cardiac disorders associated with transport  
 CC e.g. polymyositis, bradyarrhythmia, dermatomyositis, angina, neurological  
 CC disorders associated with transport e.g. amnesia, bipolar disorder,  
 CC depression, Tourette's disorder, schizophrenia, other disorders  
 CC associated with transport e.g. neurofibromatosis, sickle cell anaemia,  
 CC Wilson's disease, cataracts, infertility, hyperglycaemia, hypoglycaemia,  
 CC goitre, Cushing's disease, hypercholesterolaemia and cystinuria. Cell  
 CC proliferated disorders include cancer, actinic keratosis, cirrhosis,  
 CC arteriosclerosis, atherosclerosis, bursitis, hepatitis and psoriasis.  
 CC Neurological disorders include Alzheimer's, Pick's and Parkinson's  
 CC disease, amyotrophic lateral sclerosis, epilepsy, stroke, Huntington's  
 CC disease, multiple sclerosis, dementia and other extrapyramidal disorder,  
 CC motor neuron disorder, prion disease, metabolic disease of the nervous  
 CC system and other developmental disorders of the central nervous system,  
 CC neuromuscular disorders, metabolic, endocrine and toxic myopathies,  
 CC periodic paralysis, mental disorders including mood, anxiety; and  
 CC immunological disorders include acquired immune deficiency syndrome  
 CC (AIDS), adult respiratory distress syndrome, Addison's disease,  
 CC allergies, asthma, atherosclerosis, osteoporosis, autoimmune haemolytic  
 CC anaemia, autoimmune thyroiditis, Crohn's disease, atopic dermatitis,  
 CC Grave's disease, glomerulonephritis, rheumatoid arthritis, scleroderma,  
 CC systemic lupus erythematosus, systemic sclerosis, ulcerative colitis,  
 CC haemodialysis, urethritis, viral, bacterial, fungal, parasitic, protozoal,  
 CC helminthic infections and trauma; and muscle disorders include cardiac  
 CC myopathy, myocardiitis, polymyositis, arrhythmias and hypertension. The  
 CC TRICH polynucleotides are used in gene therapy. The present sequence is  
 CC human TRICH-3 cDNA

XX SQ Sequence 6369 BP; 1893 A; 1112 C; 1246 G; 2118 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 6369  
 Score: 8422.00 Matches: 1641  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 99.94% Mismatches: 0  
 Query Match: 99.95% Indels: 0  
 DB: 6 Gaps: 0

US-10-090-458-5 (1-1642) x AAD33648 (1-6369)

QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
 DB ATGCCACTGCAATAGGAGGAGTAGGAGTTGGAGACAGACACACTTCTACTGAAG 1297  
 QY 21 AsnTrpLeuIleLysCysArgThrLysLysSerValGlnGluLeuLeuPheProLeu 40  
 DB AATTACTTAATTAATGACAGACCAAAAGAGTAGTGTTCAGGAATCTTTTCCACTA 1357  
 QY 41 PhePheLeuPheTrpLeuIleLeuLeuSerMetMetHisProAsnLysLysTrpGluGlu 60  
 DB TTTTTTTTATTTGGTTTAAATTAATTAATAGCATGATCATCAATAAAGAAATATGAAGAA 1417  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80  
 DB GTGCCTAATATAGACTCAATCTATGGACAAGTTTACITTTCTTAATCTAATCTTTGGA 1477  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 DB TATACTCCAGTGACTAATATTAAGCAGCATCATGAGAGGAGTGTCTACTGATCATCTA 1537

QY 101 ProAspValIleIleThrGluTrpThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 DB CTGTATGTCATATTAATTAATGAAGAAATATCAAAATGAAGAAATGTTAAACATCCAGTCTC 1597  
 QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTrpGluLeuArg 140  
 DB TCTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAGAGACTCCATGCTCTATGAATCTCGT 1657  
 QY 141 PhePheProAspMetIleProValSerSerIleTrpMetAspSerArgAlaGlyCysSer 160  
 DB TTTTTCCTGATATGATCCAGTATCTTCTATATATATATGATTCAGAGCTGGCTGTTCAT 1717  
 QY 161 LysSerCysGluAlaAlaGlnTrpSerSerGlyPheThrValLeuGluAlaSerIle 180  
 DB AAATCATGTGAGGCTGCTCAGTACTGTGCTCAGGTTTACAGATTTTACAAAGCATCCATA 1777  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
 DB GATGCTGCATTTATACATTTGAGACCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACT 1837  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGlnIleAspThrPheProArgGlyVal 220  
 DB AAAGCTGTTTATATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA 1897  
 QY 221 IleLeuIleTrpLeuValIleAlaPheSerProPheGlyTrpPheLeuAlaIleHisIle 240  
 DB ATTTTAAATATACCTAGTTATAGCAATTTTACCTTTTGGCAATTCATATC 1957  
 QY 241 ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLysHisAspThr 260  
 DB GTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 2017  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTrpThrSerLeuIlePheLeuMetSerLeuLeu 280  
 DB GCCTTTTGGCTTCTCTCTATATCAAGATTTTAAATTTTCTTATGTCCTCTCTT 2077  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
 DB ATGGAGTCATTTGCGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTGATATT 2137  
 QY 301 LeuLeuPheLeuTrpGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 DB CTGCTTTTTTCTTTATGAGTATATCATCTGCTATTTTGTCTTTTAAATGCTGACACTCT 2197  
 QY 321 PheLysLysSerLysHisValGlyLysValGluPhePheValThrValAlaPheGlyPhe 340  
 DB TTTTAAAAAATCAAAACATGCGGAAATAGTTGAAATTTTGTACTGTGGCTTTTGGATT 2257  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 DB ATTTGGCTTATGATAATCTCTCATAGAAAGTTTTCCTCAAGTAGCAGCATTTGTGATATT 2317  
 QY 361 ProPheCysHisCysThrPheValIleGlyLysAlaGlnValMetHisLeuGluAspPhe 380  
 DB CCTTTCTGCTACTGTACTTTTGTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 2377  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTrpProLeuIleIleThrIle 400  
 DB AATGAAGTGTCTTCAATTTCAATTTTGAATGAGCCCATATCTCTAAATTTATTAACAAT 2437  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTrpValLeuLeuAlaValTrpLeuAspGluVal 420  
 DB ATCATGCTCACACTTAATAGTATATTTCTATGCTCTTGTGCTGTCTATCTTGTATCAAGTC 2497  
 QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTrpPheLeuLysProSerTrpTrp 440  
 DB ATTCAGGGGAATTTGGCTTACGAGCATCACTTTTATTTATTTTCTGAAGCTTCATATTGG 2557  
 QY 441 SerLysSerLysArgAsnTrpGluLeuSerGluLysAsnValAsnGlyAsnIleSer 460  
 DB TCAAGAGCAAAAGAAATTTATGAGGAGTTATCAGAGGCAATGTTAATGGAATATTAGT 2617

QY 461 PheSerGluIleLeuProValSerSerGluPheValGlyLysGluAlaIleArgile 480  
DB 2618 TTTAGTGAATTTATTGACCGAGTTCTTCAGATTTGTAGAAAAGAGCCATAGAAATT 2677  
QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
DB 2678 AGTGGTATTTCAGAGACATACAGAAAGAGGGTGAAATTTGGAGGCTTTGAGAAATTTG 2737  
QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
DB 2738 TCATTTGCATATATAGGGTCAGATTTACTGCCCTTACTTGGCCACAGTGGACAGGAAG 2797  
QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
DB 2798 AGTACATTCATGAATATTCTTTGTGGACTCTGCCACCTTCTGTATGGGTTTGGCATCTATA 2857  
QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
DB 2858 TATGGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAATGATTTGGCATT 2917  
QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
DB 2918 TGTCCACAGTTAGATATACATCTTTCATGTTTGCAGTAGAAGAAAATTTATCAATTTTG 2977  
QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
DB 2978 GCTTCAATCAAGGATACACGCAACCAATATATACAAAGAGTGCAGAGGTTTTACTA 3037  
QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
DB 3038 GATTTAGACATSCAGACTATCAAGATTAACCAAGCTAAAAAATTAAGTGTGTGCAAAA 3097  
QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
DB 3098 AGAAAGCTGTCTATTAGGAATTTGCTTCTTGGGAACCCAAAGATACTGCTCTAGATGAA 3157  
QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLeuLysTyrArg 660  
DB 3158 CCAACAGCTGGAAATGGAACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAANATACAGA 3217  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
DB 3218 AAAGCCAAATCGGGTGACAGTGTCTAGTACTCATTTTCAGATGAAGCTGACATTTCTTGA 3277  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
DB 3278 GATAGGAAGCTGTGATATACAGGAATGCTGAAATGTTGTGTTCTTCATGTTCTCTC 3337  
QY 701 LysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 3338 AAAAGTAAATGGGGATCGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACA 3397  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
DB 3398 GAATCTCTCTTCTCACTGGTTAAACACATATACCTGGAGTACTTTTATTACACAGAT 3457  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
DB 3458 GACCAACAACTTGTGTATAGTTCGCTTCAAGGACATGGACAAATTTTCAGGTTTGTGT 3517  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 3518 TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTATGGTGTTCATGACGACT 3577  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
DB 3578 TTGGAGAGCGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATTTATAGT 3637  
QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 3638 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTGTGAAATGGAA 3697  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840

DB 3698 CAGAGCTTACTTATTCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGG 3757  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
DB 3758 AAACAAACAGATGATACAAATAGCAAGTTTCATTTCTTTACCTGAAACGTAAGATAA 3817  
QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 3818 TCAGTGAGATCAGTGTGCTGCTCTTAAATTTTCTTTTTCACAGTTTCAGATTTTATGTTT 3877  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
DB 3878 TTGGTTTCATCTCTTTTAAATGCTGTGGTTCCATCAAACTTGTTCACGACTTATAT 3937  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
DB 3938 TTTCTAAAACCTGGAGACAAACCAATAAATACAAAACAAGTCTGCTTCTTCAAAATTTCT 3997  
QY 921 AlaAspSerAspIleSerAspLeuIleSerPheThrSerGlnAsnIleMetValThr 940  
DB 3998 GGTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAAGCCAGAACATAATGTGACG 4057  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
DB 4058 ATGATTTAATGACAGTGACTATGTCGCTGGCTCCCATAGTCGGCTTTAAATGTGGTG 4117  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
DB 4118 CATTCAGAAAAGGACTATGTTTTGCAGCTGTTTTCAACAGACTATGTTTATTCTTTTA 4177  
QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
DB 4178 CCTATTATTAGTGAATATCATTTAGTAACTACTATCTTTATCATTTAAATGTGAAACC 4237  
QY 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
DB 4238 ATCCAGATCTGGAGTACCCCATCTTCTCAAGAAATTTACTGATATAGTTTTTAAATTTAG 4297  
QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
DB 4298 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTTACTGTGCAATGCCACTTACIT 4357  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
DB 4358 ATGGAATAATGCAGAGAATCATTAAGATCAAGCTTATCTCAACTTAAACTTTTCAGGTC 4417  
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
DB 4418 TTGCCATCTGCATTTGGATTGACAGCTGTTGTTGATATCCCTTATTTTATCATTT 4477  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
DB 4478 CTATTATTTGATGCTAGGAAGCTTATTGGCATTTCTATTATGGAATATATTTTATATCTGTA 4537  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
DB 4538 AAGTCTCTGCTGCTGTTTTTGGCTTATTTGGTTATGTTTCCATCAGTATTCTGTTCACT 4597  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
DB 4598 TATATTGCTCTTTTCACTTTTAAAGAAAATTTTAAATACCAAGAAATTTTGGTTCATTTATC 4657  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
DB 4658 TATCTGTGGCAGCGTGGCTGTGTATGCAATCAGTCACTGAAATAACTTCTTTATGGATAC 4717  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
DB 4718 ACATTTGCAACTATTCTTCAATTATGCTTTTGTATCATCATTCCAATCTATCCACTCTTA 4777  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200

Db 4778 GGTGCTGCTGATTTCTTTTCATAGAAGATTTCTTGGAGAAGATGTACGAAAAAATGTGGACACC 4837  
QY 1201 TyrAsnProTyrAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValIleu 1220  
Db 4838 TATAATCCATGGATAGGCTTTTCAGTAGCTGTATATCGCTTACCTCCAGTGTGTACTG 4897  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGlnLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 4898 TGGATTTTCTCTTCAATATCTATCAGAAAAAATATGAGGAGCAGATCAATTAAGAAAAAT 4957  
QY 1241 ProPhePheAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 4958 CCCCTTTTTCAGAAACCTTTCAACGAAGCTTAATAATAGAAAGCTTCCAGAACCCACGAC 5017  
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 5018 AATGAGGATGAAGATGATGCTCAAGCTCAAGACTAAAGGTCAAGAGCTGATGGT 5077  
QY 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 5078 TGCCAGTGTGTGAGGAGAAACCTCCATTATGTCAGCAATTTGCTATAAAGNATATGAT 5137  
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320  
Db 5138 GACAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGCCACTAATATCATCTCT 5197  
QY 1321 PheCysValLysLysGlyGluLeuLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
Db 5198 TTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGGTCCAAATGTGTGGCAAAAGC 5257  
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 5258 ACAATTAATTAATTTCTGTGTGTGTGTATTTGACCACTTCAGCCAGGATATTTTAGGA 5317  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 5318 GATTATTTCTTCAGACAAAGTGAAGATGATGATTCCTGAAAGTGTATGGGTACTGTCT 5377  
QY 1381 GlnIleAsnProLeuTyrProAspThrThrLeuGlnGlnHisPheGluIleTyrGlyAla 1400  
Db 5378 CAGATAAACCCCTTTGTGGCCAGATACTATCATTGACGAGAACATTTTGAATTTATGGAGCT 5437  
QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db 5438 GTCAAGGAATGAGTGCAGTGACATGAAGAAGTCATAAGTCGAATAACACATGCATCT 5497  
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
Db 5498 GATTTTAAAGAACATCTTTCAGAAAGACTGTAAAGAAACTACCTGCAGGAATCAAAACGAAAG 5557  
QY 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
Db 5558 TTGTGTTTTGCTTAAGTATGCTAGGGAATCTCAGATTACTTTGCTAGATGAACCATCT 5617  
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
Db 5618 ACAGATATGGATCCCAAGCCCAACAGACACATGTGGCGAGCAATTCGAATTCGATTTAAA 5677  
QY 1481 AsnArgLysArgAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCys 1500  
Db 5678 AACAGAGCGGCTGCTATTTCTGACCACTCATTATATGAGGAGGAGCGCTGTCTGT 5737  
QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
Db 5738 GATCAGTAGCTATCATGTGTCTGGGCGAGTAAAGATGTATCGGAACAGTACACATCTA 5797  
QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
Db 5798 AAGATTAATTTGGAAAAGGCTACTTTTGGAAATTAATTAAGAGGAGCTGATAGAAAAC 5857  
QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
Db 5858 CTAGAAGTAGACCGGCTTCAAAGAGAAATTCAGTATATATTTTCCCAATGCAAGCGGTGAG 5917

QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
Db 5918 GAAAGTTTTTCTTCTATTTTGGCTTATAAATTTCTAAGAGAGATGTTTCAGTCCCTTTCA 5977  
QY 1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600  
Db 5978 CAATCTTTTTTAAAGCTGGAAGAAGCTAAACATGCTTTTGCCATTGAAGATATAGCTTT 6037  
QY 1601 SerGlnAlaThrLeuGlnGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
Db 6038 TCTCAAGCAACATTGGAACAGGTTTTTGTAGAACTCCTAAAGAACAGAGGAGGAGAT 6097  
QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyrTrpGluArgThrGlnGluAspArgVal 1640  
Db 6098 AATAGTTGTGGAACCTTTAAACAGCACACTTTGGTGGAAACGAACAACAAGATAGAGTA 6157  
QY 1641 ValPhe 1642  
Db 6158 GTATTT 6163  
RESULT 3  
ABN89594  
ID ABN89594 standard; cDNA; 6525 BP.  
XX AC ABN89594;  
XX XX 18-SEP-2002 (first entry)  
XX DE Human ATP-binding cassette transporter ABCA5 cDNA SEQ ID NO:1.  
XX KW Human; ABCA5; ABCA6; ABCA9; ABCA10; ATP-binding cassette transporter;  
KW chromosome 17; chromosome 17q; chromosome 17q24; antiarteriosclerotic;  
KW gene therapy; cholesterol; lipophilic molecule; inflammation;  
KW prostaglandin; prostacyclin; arteriosclerosis; transport; gene; ss.  
XX OS Homo sapiens.  
XX WO200246458-A2.  
XX XX 13-JUN-2002.  
XX PF 07-DEC-2001; 2001WO-BP015401.  
XX PR 07-DEC-2000; 2000EP-00403440.  
XX PR 23-JAN-2001; 2001US-0263231P.  
XX XX (AVET ) AVENTIS PHARMA SA.  
XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX PI Denefle P, Rosier-Montus M, Prades C, Arnould-Reguigne I;  
PI Duverger N, Allikmets R, Dean M;  
XX XX WPI; 2002-557584/59.  
DR P-PSDB; AB881574.  
XX XX A novel nucleic acid corresponding to ATP-binding cassette transporter  
PT genes and the encoded polypeptide, useful for preventing or treating a  
PT dysfunction in reverse transport of cholesterol.  
XX PS Claim 1; Page 151-153; 216pp; English.  
XX CC The present invention describes human ATP-binding cassette transporters  
CC (ABC). Specifically described are the human ABCA5, ABCA6, ABCA9 and  
CC ABCA10 genes (see ABN89594 to ABN89597) which encode the proteins given  
CC in AB881574 to AB881577. ABN89598 to ABN89715 represent ABCA5, ABCA6,  
CC ABCA9 and ABCA10 nucleotide fragments; and ABN89716 to ABN89806 represent  
CC primers for ABCA5, ABCA6, ABCA9 and ABCA10 genes which are used in the  
CC exemplification of the present invention. The ABC sequences have  
CC antiarteriosclerotic activities and can be used in gene therapy. ABC  
CC sequences can be used in the manufacture of a medicament intended for the  
CC prevention and/or treatment of a subject affected by a dysfunction in the  
CC reverse transport of cholesterol. The ABC proteins are involved in the



CC reverse transport of cholesterol, in membrane transport of lipophilic  
 CC molecules, in particular inflammation mediating substance such as  
 CC prostaglandins and prostacyclins, or in any pathology whose candidate  
 CC chromosomal region is situated on chromosome 17. They are also useful for  
 CC the manufacture of a medicament intended for prevention of  
 CC arteriosclerosis in various forms. The ABCA5, ABCA6, ABCA9 and ABCA10  
 CC genes are located to chromosome 17, more specifically to the 17q24 locus  
 XX  
 SQ Sequence 6525 BP; 1973 A; 1108 C; 1258 G; 2185 T; 0 U; 1 Other;

## Alignment Scores:

Pred. No.: 0 Length: 6525  
 Score: 8422.00 Matches: 1641  
 Percent Similarity: 99.94% Conservative: 0  
 Best Local Similarity: 99.94% Mismatches: 1  
 Query Match: 99.95% Indels: 0  
 DB: 6 Gaps: 0

US-10-090-458-5 (1-1642) x ABN89594 (1-6525)

QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
 Db 1011 ATGTCCACTGCATTAAGGAGGTAGGAGTTGGAGACAGACAGAACACTTCTACTGAAG 1070  
 QY 21 AsnTyrLeuLeuLysCysArgThrLysLysSerValGlnGluLeuLeuPheProLeu 40  
 Db 1071 AATTACTTAAATTAATGCAGAACCAAAAGAGTAGTGTTCAGGAAATTCCTTTCCACTA 1130  
 QY 41 PhePheLeuPheTrpLeuLeuLeuSerMetMetHISProAsnLysLysTyrGluGlu 60  
 Db 1131 TTTTITTTTATTTTGGTTAATAATTAATAGCATGATGCATCCAAATAAGAAATATGAAGAA 1190  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80  
 Db 1191 GTGCCTAATATAGAACTCAATCCTATGACAGAGTTTACTCTTCTAATCTTCTTGA 1250  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 Db 1251 TATACTCCAGTCACTAATATTACAGCAGCATCATGCAGAAAGTCTTACTGATCATCTA 1310  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 Db 1311 CCTGATGTCATTAATTACTGAAGAATATCAAAATGAAGAAAGAAATGTTAACATCCAGTCTC 1370  
 QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
 Db 1371 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAGACTCCATGCTCTATGAATTCGT 1430  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 Db 1431 TTTTITTCCTGATATGATCCAGTATCTTCTATTATATGGATTCAAGAGCTGGCTGTTC 1490  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 Db 1491 AAATCATGTAGGCTGCTCAGTACTGTGCTCCTCAGTTTCACAGTTTTCACAGCATCCATA 1550  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
 Db 1551 GATGCTGCCATTATACAGTTGAAGCAATGTTTCTCTTGAAGAGAGCTGGAGTCAACT 1610  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
 Db 1611 AAAGCTGTTATTATGGAGAAATCTGCTGTTGAGAAATAGATACCTTTCCCGGAGGAGTA 1670  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 Db 1671 AATTTAATATACCTAGTATTAGCAATTTTCCACCTTTTGGATACCTTTTGGCAATTCATATC 1730  
 QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 Db 1731 GTAGCAGAAAAAAGAAAAAATAAAGAAATTTTAAAGATTAATGGGACTTTTCATGATACT 1790  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280

Db 1791 GCCTTTGGCTTCTCGGTCTCTCTATATACAACTTTAAATTTTCTTATGTCCTCTCT 1850  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
 Db 1851 ATGGCAGTCATTGGACAGCTTCTTTGTTATTCTCTCAAGTAGCAGCATTTGTGATATT 1910  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 Db 1911 CTGCTTTTTCCTTTATGGATTATCATCTGTATTATTTTGTCTTAAATGTCACACTCTT 1970  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 Db 1971 TTTTAAAAATCAACATAGTGGAAATAGTTGAAATTTTGTCTGTGGCTTTTGGATTT 2030  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 Db 2031 ATTGGCTTATGATAATCTCATAGAAAGTTTTCCAAATCGTTAGTGTGGCTTTTCAGT 2090  
 QY 361 PropheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 Db 2091 CCTTCTGTCAGTACTTTTGTGATTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2150  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 Db 2151 AATGAAGGTGCTTCAATTTTCAATTTGACTGCAGGCCCATATCCTCTAATTATTACAAAT 2210  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 Db 2211 ATCATCTCACACTTAAATAGTATATCTATGTCCTTGGCTGTCTATCTTGTATCAAGTC 2270  
 QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 Db 2271 ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTATATATTTTCTGAAGCTTCATATTGG 2330  
 QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 Db 2331 TCAAGAGTAAAGAAATATGAGGAGTTATCAGAGGGCAATGTTAATGGAATATTAGT 2390  
 QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
 Db 2391 TTTAGTGAATATTATGAGCCAGTTTCTCAGAAATTTAGGAAAGAACCCATAGAAT 2450  
 QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyLysAsnValGluAlaLeuArgAsnLeu 500  
 Db 2451 AGTGGTATTCAAGAGACATACAGAAAGAGGTGAAATGTGGAGCTTTTGAGAAATTTG 2510  
 QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuGlyHisSerGlyThrGlyLys 520  
 Db 2511 TCATTTGACATATATGAGGTCAGATTACTGCTTACTTGGCCACAGTGGACAGGAAG 2570  
 QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540  
 Db 2571 AGTACATTGATGAATATTCTTTGTGACTCTGCCACCTTCTGATGGGTTCATCTATA 2630  
 QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 Db 2631 TATGGACACAGAGCTCTCAGAAATAGTAAATGTTTGAAGCAGAAATAATGATTGGCAT 2690  
 QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 Db 2691 TGTCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAAGAAATTTATCAATTTTG 2750  
 QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 Db 2751 GCTTCAATCAAGGGATACAGCCAAATAATAACAAGAGTGGAGAGGTTTACTTA 2810  
 QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 Db 2811 GATTTAGACATGCAGACTATCAAGATATCAAGCTAAAGCTAAAGAAATTAAGTGTGTGCTCAAAA 2870  
 QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640

Db 2871 AGAAGCTGTCATTAGGAATTCGTCTTCTGGGAACCCAAAGATACCTGCTGCTAGATGAA 2930  
Qy 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
Db 2931 CCAACAGCTGGAATGGACCCCTGTTCTCGACATATTCGTATGGAATCTTTTAAATACAGA 2990  
Qy 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 2991 AAAGCCAAATCGGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGTCACATCTCTGCA 3050  
Qy 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 3051 GATAGGAAGCTGTGATATCACAGGAATGCTGAATGTGTGCTTCTTCAATGTTCCTC 3110  
Qy 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 3111 AAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACA 3170  
Qy 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db 3171 GAATCTCTTCTTCACTGGTTAAACAACATATACCTGGAGCTACTTTATTACAACAGAA 3230  
Qy 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 3231 GACCAACAACATGTGTATAGTTCGCTTCAAGGACATGGACAAATTTTCAGGTTCGTTT 3290  
Qy 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 3291 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCATTTCTTATGGTGTTCATGACGACT 3350  
Qy 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 3351 TTGGGAAGACGTATTTTAAAGCTAGAGCTTGAAGCAGAAATTTGACCAAGCAGATATATAGT 3410  
Qy 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 3411 GTATTTACTCGACGCCACTCGAGGAAGAAATGATTCAAAATCTTTTGATGAATGGAA 3470  
Qy 821 GlnSerLeuLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 3471 CAGAGCTTACTATTCCTTCTGAAACCAAGCGCTTCTAGTGAGCACCATGAGCCTTTGG 3530  
Qy 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
Db 3531 AAACAACAGATGTATACATAGCAAGTTTCATTTCTTTACCTTGAACACGCTGAAGTAA 3590  
Qy 861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 3591 TCAGTGAGATCAGTGTTGCTTCTGCTTTTAATTTTATTTTTCACAGTTCAGATTTTATGTTT 3650  
Qy 881 LeuValHisHisSerPheLysAsnAlaValValProLysLysLeuValProAspLeuTyr 900  
Db 3651 TTGGTTTCATCACTCTTTTAAAAATGCTGTGGTTCCCATCAAACTTGTTCAGACCTTATAT 3710  
Qy 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 3711 TTTCATAAACCTGGAGACAAACCATAAATACAAACCAAGCTGCTTCTTCAAAAATCT 3770  
Qy 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 3771 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAAGCAGCAATATGCTGACG 3830  
Qy 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 3831 ATGATTAAATGACGTAGCTATGTATCCGTGGCTCCCATAGTGGCGCTTTTAAATGTGATG 3890  
Qy 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 3891 CATTCAGAAAGACATATGTTTTCGACGTGTTTTCACAGTACTATGTTTATTCITTA 3950  
Qy 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 3951 CCTATATTAGTGAATATCATTTAGTAACACTACTATCTTTATCTTTTAAATGTGACTGAAACC 4010

Qy 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 4011 ATCCAGATCTGAGTACCCCATCTTCTTCAAGAAATTAAGTATATATTTTAAAAATTGAG 4070  
Qy 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 4071 CTGTATTTTCAAGCAGCTTGTCTTGGAAATCATTTGTACTGCATGCCACCTTACTTTGCC 4130  
Qy 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 4131 ATGGAAAAATGCAGAGAAATCAAGATCAAGCTTATATCTCAACTTAAACTTTTCAGGCTT 4190  
Qy 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleIle 1080  
Db 4191 TTGCCATCTGCATATTTGATTTGGACACAGCTGTTGTGTATATCCCTTATTTTATCAT 4250  
Qy 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 4251 CTATTTTGTATGCTAGGAAGCTTACTGGCAATTTCAATATGATATATTTTATATCTGTA 4310  
Qy 1101 LysPheLeuAlaValValPheCysIleuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 4311 AAGTTCTCTGCTGTGGTGTTCCTTATTTGGTATGTTCCATCACTTATTTCTGTTCACT 4370  
Qy 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 4371 TATATTGCTTCTTTCACCTTTAAGAAAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 4430  
Qy 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 4431 TATCTGTGGCAGCTTGNCTTGTATTGCAATCACTGAAATAAATCTTCTTATGGGATAC 4490  
Qy 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 4491 ACBAATGCACTATTTCTTCAATTTGCTTTGTATCACTCAATTCATCTTCACTTCTTA 4550  
Qy 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
Db 4551 GGTGCTGCTGATTTCTTTCATAAAGATTTCTTGGAAAGAAATGTACGAAAAAATGTGGACAC 4610  
Qy 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 4611 TATAATCCATGGGATAGGCTTTCAGTAGCTGTATATGCTTATGCTTACTGTCAGTGTGCTG 4670  
Qy 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 4671 TGGATTTTCTCTTCAATCTATCGAAAAAATATGGAGGCAGATCAATAAGAAAAAGAT 4730  
Qy 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAsp 1260  
Db 4731 CCTTTTTCAGAAACCTTTCAACGAAGTCTAAAAATAGGAAGCTTCCAGAACCCACAGAC 4790  
Qy 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 4791 AATGAGATGAAGATGAAGATGCTCAAGCTGAAGACTAAAGGTCAAGAGCTGATGGGT 4850  
Qy 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 4851 TGCAGTGTGTGTGAGGAGAACCATTCATTTATGGTCAGCAATTTGCATAAAGAAATATGAT 4910  
Qy 1301 AspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer 1320  
Db 4911 GACAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGCACAACTAAATACATCTCT 4970  
Qy 1321 PheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
Db 4971 TTCTGTGTGAAAAAAGAGAGATCTTAGGACTATTGGGTCCAAATGGTGTCTGGCAAAAGC 5030  
Qy 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 5031 ACATTTTAAATATTTCTGTTGGTGTGATTTGAACCAACTTCAGGCCAGGTATTTTATGGA 5090

```
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLeuLysCysMetGlyTyrCysPro 1380
DB 5091 GATTATTCTTCAGAGACAAGTGAAGATGATGATTCACCTCAAGTGTATGGGTACTGTGCT 5150
QY 1381 GlnIleAenProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400
DB 5151 CAGATAAACCTTTGTGGCCAGATACATCATGTCAGGACATTTTGAATTTATGGAGCT 5210
QY 1401 ValIysGlyMetSerAlaSerAspMetIysGluValIleSerArgIleThrHisAlaLeu 1420
DB 5211 GTCAAGGAATGAGTGAAGTGAATGAAGAGTCAATAGTTCGAATTAACACATGCACCT 5270
QY 1421 AspLeuLysGluHisLeuGlnLysThrValIysLysLeuProAlaGlyIleLysArgLys 1440
DB 5271 GATTTAAGAAGACATCTTCAGAGACTGTAAAGAACTACTCTGCAGAGATCAACAGAAAG 5330
QY 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460
DB 5331 TTGTGTTTTGCTCTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCATCT 5390
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLys 1480
DB 5391 ACAGTATGGATCCCAAGCCCAACAGCACATGTGGCGAGCATTCGAATCGATTGATTA 5450
QY 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500
DB 5451 AACAGAAAGCGGCTGCTATTCTGACCACTCATTATATGAGGAGGAGGAGGCTGTCTGT 5510
QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520
DB 5511 GATCGAGTAGCTATCATGTGCTGGGCGATTAGATGTATCGGAACAGTACACATCTA 5570
QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540
DB 5571 AAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTTAAATTTGAAGGACTGGATAGAAAC 5630
QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560
DB 5631 CTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCCGTCAG 5690
QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580
DB 5691 GAAAGTTTTCTTCTATTTTGGCTATTAAATTTCTTAAGGAGATGTTCACTCCCTTCA 5750
QY 1581 GlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluTyrSerPhe 1600
DB 5751 CAATCTTTTTTAAAGCTGAAGAACTAAACATGCTTTTGCCATTGAAGATATAGCTTT 5810
QY 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620
DB 5811 TCTCAAGCAACATTTGGAACAGGTTTTTTGAGAACTCACTAAAGAACAGAGGAGAGAT 5870
QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyrTrpGluArgThrGlnGluAspArgVal 1640
DB 5871 AATAGTCTGGAACCTTTAAACAGCACACTTTGGTGGGAACCAACACAGAAGATAGATA 5930
QY 1641 ValPhe 1642
DB 5931 GPATTT 5936
RESULT 4
AAD37618
ID AAD37618 standard; cDNA; 4929 BP.
XX AAD37618;
AC AAD37618;
XX
XX 10-SEP-2002 (first entry)
DT
XX
XX Human transporter protein cDNA #1.
DE
XX Human; novel human protein; NHP; transporter protein; mental disorder;
KW cancer; gene therapy; drug screening; nutraceutical application; gene;
KW cosmetic application; polymorphism; ss.
```

```
XX Homo sapiens.
OS
XX
FH Key
CDS
Location/Qualifiers
1..4929
/*tag= a
/product= "Human transporter protein #1"
/transl_except= (pos:808..810, aa:Tyx)
/transl_except= (pos:2494..2496, aa:Ser)
/transl_except= (pos:2878..2880, aa:Val)
810
misc_feature
/*tag= b
/note= "This degenerate base represents a polymorphic
site"
2494
misc_feature
/*tag= c
/note= "This degenerate base represents a polymorphic
site"
2878
misc_feature
/*tag= d
/note= "This degenerate base represents a polymorphic
site"
310
WO200231147-A2.
18-APR-2002.
04-OCT-2001; 2001WO-US031113.
10-OCT-2000; 2000US-0239629P.
(LEXI-) LEXICON GENETICS INC.
Hu Y, Nepomnichy B;
WPI: 2002-454552/48.
P-PADB; AA233656.
Novel nucleic acid molecule encoding novel human proteins, useful for
therapeutic, diagnostic and pharmacogenomic applications.
Claim 1; Page 34-35; 46pp; English.
The present sequence is a cDNA coding for novel human protein (NHP),
human transporter protein. NHPs shares structural similarity with the
mammalian ATP-binding cassette (ABC) transporters and multidrug
resistance transporters. NHP polynucleotides are useful for the
therapeutic, diagnostic and pharmacogenomic applications. They are used
for detecting and treating mental disorders and cancers. They are also
used in gene therapy. NHP polypeptides are useful for diagnosis, drug
screening, clinical trial monitoring, treatment of diseases and
disorders, and cosmetic or nutraceutical applications
Sequence 4929 BP; 1544 A; 827 C; 950 G; 1604 T; 0 U; 4 Other;
Alignment Scores:
Pred. No.: 0 Length: 4929
Score: 8395.00 Matches: 1637
Percent Similarity: 93.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 5
Query Match: 99.63% Indels: 0
DB: 6 Gaps: 0
US-10-090-458-5 (1-1642) x AAD37618 (1-4929)
QY 1 MetSerThrAlaIleArgGluValGlyValTyrArgGlnThrArgThrLeuLeuLys 20
DB 1 ATGCCTACCTGCAATTAGGGAGGTAGAGTTTGGACAGACAGACACATCTTACTGAG 60
QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnIleLeuPheProLeu 40
DB 61 AATTACTTAAATTAATGCAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 120
```

QY 41 PhePheLeuPheThrLeuIleLeuIleSerMetMetHisProAsnLysLysTyGluGlu 60  
 Db 121 TTTTITTTTATTTTGGTTAATATTAATTAGCATGATCCAAATAAGAAATATGAAGA 180  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
 Db 181 GTGCCTAATATAGAACTCAATCCTATGAGCAAGTTTACTCTTCTTAATCTAATCTTGG 240  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 Db 241 TATACTCCAGTGACTAATATTAACAGCAGCATCTGCAGAAAGTGTCTACTGATCATCTA 300  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 Db 301 CCTGATGTCATAATTAATCAAAATGAAAGAAATGTTAAACATCCAGTCTC 360  
 QY 121 SerLysProSerAsnPheValGlyValPheLysAspSerMetSerTyrGluLeuArg 140  
 Db 361 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAGACTCCATGCTCATATGAACCTTCTGT 420  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 Db 421 TTTTITTCCTGATGATGATCCAGTACTTCTATTTATATGGATTCAAGAGCTGGCTGTCA 480  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 Db 481 AAATCATGTGAGGCTGCTCAGTACTGCTCAGGCTTTCAGAGTTTCAAGCATCCATA 540  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
 Db 541 GATGCTGCCATTATACAGTTGAGACCAATGTTTCTCTTTGGAAGGAGCTGGAGTCAACT 600  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
 Db 601 AAAGCTGTATTATGGGAAACTGCTGTGTAGAAATAGATACCTTTCCCGCAGGAGTA 660  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 Db 661 ATTTTAATATACCTAGTTATAGCATTTTCACTTTTGGATCTTTTGGCAATTCATATC 720  
 QY 241 ValAlaGluLysGluLysLysLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 Db 721 GTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 780  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 Db 781 GCCTTTTGGCTTCTCTGGGTTCTCTATAWACAAGTTTAAATTTTCTTATGTCCTCTCT 840  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300  
 Db 841 ATGGCAGTCATTCGACAGCTTCTTTGTTATTTCTCCTCAAGTAGCAGCAATTTGTATTT 900  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 Db 901 CTGCTTTTCTCTTTATGGATATCATCTGTATTTTGTCTTTAATGCTGACACCTCTT 960  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 Db 961 TTTTAAAAATCAAAACATGCGGATAGTTGAAATTTTGTGTTTGTGCTTTTGGATTT 1020  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 Db 1021 ATTTGGCTTATGATAATCCTCATGAAAGTTTCCCAATCGTTAGTGTGGCTTTTTCAGT 1080  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 Db 1081 CTTTCTGTCTACTGTACTTTTGTGATGTTGTTATGACAGGCTCATGCAATTTAGAGATTTT 1140  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 Db 1141 AATGAAGGTGCTTCAATTTCAAAATTTGACTGCAGGCCCATATCTCTAAATTTATCAAT 1200  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420

Db 1201 ATCATGCTCACCTTAATAGTATTTCTATGTCTCTCTTGGCTGTCTATCTTGTATCAAGTC 1260  
 QY 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 Db 1261 ATTCCAGGGAAATTTGGCTTACGGAGATCATCTTTATATATTTTCTGAAGCCTTCATATTGG 1320  
 QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 Db 1321 TCAAAGAGCAAAAGAAATTTATGAGGAGTTATCAGGGCAATGTTAATGGAATATTAGT 1380  
 QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
 Db 1381 TTTAGTGAATATTATGAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCAATGAAGATT 1440  
 QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
 Db 1441 AGTGGTATTTCAGAAAGACATACAGAAAGAGGGTGAATGTGGAGCTTTGAGAAATTTG 1500  
 QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 Db 1501 TCATTTGACATATATGAGGTCAGATTACTGCTTACTTTGGCCACAGTGGACAGGAAG 1560  
 QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerSerAspGlyPheAlaSerIle 540  
 Db 1561 AGTACATTTGATGAATATTCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1620  
 QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 Db 1621 TATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATTGSCATT 1680  
 QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 Db 1681 TGTCCACAGTTAGATATACACTTTGATGTTTGTACAGTAGAAGAAATTTATCAATTTTG 1740  
 QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 Db 1741 GCTTCAATCAAGGATATACAGCCCAATATATATCAAGAGAGTGACAGAGGTTTACTA 1800  
 QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 Db 1801 GATTTAGACATGCAGACTATCAAGATAACCAAGCTTAAAGAAATTAAGTGTGTGTCAAAA 1860  
 QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysLleLeuLeuLeuAspGlu 640  
 Db 1861 AGAAAGCTGTATAGAAATGCTGTTCTTGGAGACCCCAAGATATCTGCTGTAGTAGAA 1920  
 QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyrArg 660  
 Db 1921 CCAACAGCTGGAATGGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAATACAGA 1980  
 QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 Db 1981 AAGCCAATCGGGTGACGTGTTCACTACTCAATTCATGATGAGAGCTGACATTTCTTGA 2040  
 QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 Db 2041 GATAGAAAGCTGTATATCAAGGAATGCTGAAATGTTGTTGTTCTTCAATGTTCTCTC 2100  
 QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 Db 2101 AAAAGTAAATCGGGGATCGGCTACCGCTGAGCATGTATCATAGACAAATATTGTGCGACA 2160  
 QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyValAlaThrLeuLeuGlnAsn 740  
 Db 2161 GAATCTCTTCTTCTCTGTTAAACAAATATATCTGGAGCTACTTTATTACACAGAAAT 2220  
 QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
 Db 2221 GACCAACAACTGTGTATAGTGTGCTTTCAAGGACATGCAGAAATTTTTCAGGTTGTTT 2280  
 QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780

Db 2281 TCTGCCCTAGACAGTCATTCAAAATTTGGTGGCCATTTCTTATGGGGTTTCCATGACGACT 2340  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2341 TTGGAAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATTATAGT 2400  
QY 801 ValPheThrGlnGlnProLeuGluGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2401 GTATTACTCAGCAGCCACTGGAGAGAAATGGATTCAAAATCTTTTGATGAATGGAA 2460  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
Db 2461 CAGACCTTACTTATTTCTTCTGAAACCAAGCTKCTCTAGTGAGCACCAGAGCCTTGG 2520  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
Db 2521 AAACAACAGATGTATACAAATAGCAAAAGTTTCATTTCTTACCTTGAAACGTGAAGTAAA 2580  
QY 861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2581 TCAGTGAGATCAGTGTCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTT 2640  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 2641 TTGGTTTCATCACTCTTTTAAAAATGCTGTGGTTCCTCAATCAAACTTGTTCCAGACTTATAT 2700  
QY 901 PheLeuLysProGluAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 2701 TTTCFAAAACCTGGAGACAAACCATAAATACAAACACAGCTGCTCTTCTTCAAAATCT 2760  
QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 2761 GCTGACTCAGATATCAGTGAICTTATTAGCTTTTTCACAGCCAGAACATAAATGCTGACG 2820  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 2821 ATGATTAATAGCAGTACTATGTATCCGTGGCTCCCATAGTGGCGCTTTTAAATGTGRTG 2880  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 2881 CATTCAGAAAGACATGTTTTCACAGCTGTTTCACAGACTATAGTTTATTCCTTA 2940  
QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 2941 CCTATATTAGTCAGATATCATTTAGTAACTACTATCTTTATCATTTTAAATGTGACTGAAACC 3000  
QY 1001 IleGlnIleTyrSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 3001 ATCCAGATCTGGAGTACCCCATCTTTTCAGAAATTTACTGATATAGTTTAAATTTGAG 3060  
QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGluIleIleValThrAlaMetProProTyrPheAla 1040  
Db 3061 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTACTGCAATGCCACCTTACTTTGCC 3120  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3121 ATGGAAATTCAGAGAAATCATTAAGATCAAAAGCTTAACTCACTTAAACTTTTCAGTCTT 3180  
QY 1061 LeuProSerAlaTyrTyrIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
Db 3181 TTGCCATCTGCATATTGGATTGGACAAGCTGTGTGTGATATCCCTTATTTTATCATT 3240  
QY 1081 LeuIleLeuMetLeuGlySerLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3241 CTATATTGATGCTAGGAAGCTTATGGCAATTCATTTATGGATTATATTTTATACGTGA 3300  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 3301 AAGTTCCTTGCTGTGCTTTTTCCTTATTTGCTTATGTTATGTCATCAGTTATTTCTGTCAC 3360  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTyrSerPheIle 1140  
Db 3361 TATATTGCTCTTCTTACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420

QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3421 TATTTCTGTGGCAGCGTTGGCTTGTATTGCAATCACTGAAATAAATCTTCTTTATGGGATAC 3480  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 3481 ACAATTTGCAACATTTCTTCAATATGCCCTTTTGTATCATTCATTTCCAATCTATCCACTTTCA 3540  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTyrLysAsnValArgLysAsnValAspThr 1200  
Db 3541 GGTGGCTTGATTTCTTTCATAAAGATTTCTTGGAAAGATGTACGAAAAAATGTGCACACC 3600  
QY 1201 TyrAsnProTyrAspArgLeuSerValAlaValIleIleSerProTyrLeuGlnCysValLeu 1220  
Db 3601 TATAATCCATGGATAGGCTTTTCAGTAGCTGTATTATCGCCTTACCTCGAGTGTACTG 3660  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyValArgSerIleArgLysAsp 1240  
Db 3661 TGGATTTTCTCTTACAACTACTATGAGAAAAAATATGGAGGCAGATCAATAAGAAAAAGAT 3720  
QY 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAsp 1260  
Db 3721 CCGTTTTTTCAGAAACCTTTTCAACGAAGTCTAAAAATAGGAAGCTTCCAGAACCCACAGAC 3780  
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 3781 NATGAGATGAAGATGAAGATGTCAAAGCTGAAGACTTAAAGGTCAAAGAGCTGATGGT 3840  
QY 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 3841 TGGCAGTGTGTGAGGAGAAACCATCTATATGGTCAAGCAATTTTCATAAAGAAATATGAT 3900  
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320  
Db 3901 GACAGAAAGATTTTCTTCTTTCAGAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 3960  
QY 1321 PheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
Db 3961 TTCTGTGTGAAAAAGAGAGAGATCTTAGGACTATTGGGTCCAAATGGTCTCGCAAAAGC 4020  
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 4021 ACAATTTATTAATTTCTGGTTGGTGATTTGAACCAACTTCAGGCCAGGTATTTTAGGA 4080  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 4081 GATTTATTTCTCAGAGACAAAGTGAAGATGATGATTCATCTGAAGTGTATGGGTACTGTCTCT 4140  
QY 1381 GlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
Db 4141 CAGATAAACCCCTTTGTGGCCAGATATCATTTGCAGGAACATTTTGAATTTATGGAGCT 4200  
QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db 4201 GTCAAGGAATGAGTGCACAGTCAAGTGAAGAGTCAATAGTCAATAACACATGCACTT 4260  
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
Db 4261 GATTTAAAGAAACATCTTCAGAGAGCTGAAAGAACTACCTCGCAGGAATCAAAACGAAG 4320  
QY 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
Db 4321 TTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCATCT 4380  
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLys 1480  
Db 4381 ACAGGTATGGATCCCAAGCCAAACAGACATGTGGCGAGCAATTCGAATTCGATTTAAA 4440  
QY 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500  
Db 4441 AACAGAAAGCGGCTGCTATTCTTGACCACTCATATATGGAGGAGGAGGAGGCTGTCTGT 4500

QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
Db 4501 GATCAGTAGCTATCATGTGCTGGCAGTTAAGATGTATCGGAACAGTACAACATCTA 4560  
QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
Db 4561 AAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAAC 4620  
QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
Db 4621 CTAGAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTCCCAAAATGCAAGCGGTGAG 4680  
QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
Db 4681 GAAAGTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTCA 4740  
QY 1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGlyTyrSerPhe 1600  
Db 4741 CAATCTTTTAAAGCTGGAAGACTAAACATGCTTTTGGCAATTTGAAGAATATAGCTTT 4800  
QY 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
Db 4801 TCTCAAGCAACATTGGACAGGTTTGTAGACTCCTAAAGAACAGAGGAGAT 4860  
QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluA-gThrGlnGluAspArgVal 1640  
Db 4861 AATAGTTGTGGAACCTTAAACAGCACACTTTGGTGGGAACCAACACAGAAGATAGATA 4920  
QY 1641 ValPhe 1642  
Db 4921 GTATTT 4926

## RESULT 5

AAD37620  
ID AAD37620 standard; cDNA; 5262 BP.

XX AC AAD37620;

XX DT 10-SEP-2002 (first entry)

XX DE Human transporter protein cDNA #3.

XX KW Human; novel human protein; NHP; transporter protein; mental disorder;  
cancer; gene therapy; drug screening; nutritional application;  
cosmetic application; polymorphism; ss.

XX OS Homo sapiens.

XX PN WO200231147-A2.

XX PD 18-APR-2002.

XX PF 04-OCT-2001; 2001WO-US031113.

XX PR 10-OCT-2000; 2000US-0239629P.

XX PA (LEXI-) LEXICON GENETICS INC.

XX PI Hu Y, Nepomnichy B;

XX DR WPI; 2002-454552/48.

XX PT Novel nucleic acid molecule encoding novel human proteins, useful for  
therapeutic, diagnostic and pharmacogenomic applications.

XX PS Disclosure; Page 45-46; 46pp; English.

XX CC The present sequence is a cDNA coding for novel human protein (NHP),  
human transporter protein. NHPs share structural similarity with the  
mammalian ATP-binding cassette (ABC) transporters and multidrug  
resistance transporters. NHP polynucleotides are useful for the  
therapeutic, diagnostic and pharmacogenomic applications. They are used  
for detecting and treating mental disorders and cancers. They are also

CC used in gene therapy. NHP polypeptides are useful for diagnosis, drug  
screening, clinical trial monitoring, treatment of diseases and  
CC disorders, and cosmetic or nutraceutical applications  
XX  
SQ Sequence 5262 BP; 1615 A; 888 C; 1012 G; 1743 T; 0 U; 4 Other;

## Alignment Scores:

Pred. No.: 0 Length: 5262  
Score: 8395.00 Matches: 1637  
Percent Similarity: 99.70% Conservative: 0  
Best Local Similarity: 99.70% Mismatches: 5  
Query Match: 99.63% Indels: 0  
DB: 6 Gaps: 0

US-10-090-458-5 (1-1642) x AAD37620 (1-5262)

QY 1 MetSerThrAlaIleArgGluValTrpArgGlnThrArgThrLeuLeuLeu 20  
Db 255 ATGTCACCTGCATATTAGGAGGTAGAGTTTGGAGACAGACCAACACTTCTACTGAAG 314  
QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGluIleLeuPheProLeu 40  
Db 315 AATTACTTAATTAATGCAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 374  
QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
Db 375 TTTTCTTTTATTTGGTAAATTAATTAGCATGATCATCCAAATAAAGAAATATGAAGAA 434  
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80  
Db 435 GTGCTTAATATAGAACTCAATCTATGGACAGTTTACTCTTTCTTAATCTAATTTCTTGA 494  
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
Db 495 TATACTCCAGTGACTAATATTACAGCAGCATCATGCAGAAAGTGTTCTACTGATCATCTA 554  
QY 101 ProAspValIleIleThrGluGlyTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
Db 555 CCTGATGTCAATAATTACTGAAGAAATATACAATGAAAGAAATGTTACATCCAGTCTC 614  
QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
Db 615 TCTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAAGACTCCATGTCTCTATGAATCGT 674  
QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
Db 675 TTTTCTTCTGATATGATTCAGTATCTTCTATTATATGATTCAGAGCTGGCTCTTCA 734  
QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
Db 735 AATCATGTGAGGCTGCTCAGTACTGTCTCAGGTTTTCACAGTTTACAAGCATCCATA 794  
QY 181 AspAlaAlaIleIleGlnLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
Db 795 GATGCTGCCATTATACAGTTGAGACCAATGTTCTCTTTTGGAGAGAGTGGAGTCACT 854  
QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
Db 855 AAGCTGTTTATATGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA 914  
QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
Db 915 ATTTTAATATACCTAGTATTAGCATTTTCACTTTTGGTACTTTTGGCAATTCATATC 974  
QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
Db 975 GTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATATGGGACTTCATGATACT 1034  
QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
Db 1035 GCCTTTTGGCTTCCCTGGGTCTTCTATAWACAAGTTTAATTTTCTTATGTCCTCTT 1094  
QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300

1095 ATGGCAGTCATTGCGACAGCTCTTTGTTATTTCTCTCAAGTAGCAGCAATTTGATATTT 1154  
301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
1155 CTGCTTTTTCCTTTATGGATTATCATCTGTAATTTTGTCTTAATGCTGACACCTCTT 1214  
321 PheLysLysSerLysHisValGlyLeuValGluPhePheValThrValAlaPheGlyPhe 340  
1215 TTTAAAAAATCAAAACATGGGAATAGTTGAAATTTTGTATCTGTGGCTTTTGGATTT 1274  
341 IleGlyLeuMetIleLeuIleGluSerPheProLysSerLeuValTrrPhePheSer 360  
1275 ATTGGCCTTATGATAATCTCATGAAAGTTTCCCAATCGTAGTGGCTTTTCAGT 1334  
361 ProPheCysHisCysThrPheValIleGlyLeuAlaGlnValMetHisLeuGluAspPhe 380  
1335 CCTTCTCTGCTACGTACTTTTGTGATTGGTATTTGACAGGTCATGCATTTAGAAGATTT 1394  
381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
1395 AATGAAGGTGCTTTCATTTTCAAAATTGACTGCAGGCCCATATCTCTAAATTTATTACAAT 1454  
401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
1455 ATCATGCTCACACTTAATAGTATATTCATGTCTCTTGGCTGTCTATCTTTGATCAAGTC 1514  
421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
1515 ATTCCAGGGGAATTTGGCTTACGGAGATCACTTTATATTTTCTGAAGCCTTCATATTGG 1574  
441 SerLysSerLysArgAsnTyrGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
1575 TCAAGAGCAAAAGAAATATGAGAGATTATCAGAGGCCAATGTAAATGGAAATATTAGT 1634  
461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
1635 TTTAGTGAATTTTGGCCGATTTCTTCAGATTTGTAGGAAAGAACCATTAAGAAAT 1694  
481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
1695 AGTGGTATTTCAGAGACATACAGAAAGAGGGTGAAATGTGGAGGCTTTTGAGAAATTTG 1754  
501 SerPheAspIleTyrGlyGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
1755 TCATTTGACATATATGAGGGTCAGATTACTGCTTACTTTGCCACAGTGGACAGGAAG 1814  
521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
1815 AGTACATTGATGATATTTCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1874  
541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
1875 TATGGACACAGAGTCTCAGAAATAGATGAATTTTGAAGCAAGAAATGATTTGGCAT 1934  
561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
1935 TGTCACAGCTTAGATATACATTTGATGTTTGTAGTGGTGAAGAAATTTATCAATTTTG 1994  
581 AlaSerIleGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
1995 GCTTCAATCAAGGGGATACCAACCCCAATATAAACAAGAGTGCAGAGGTTTACTA 2054  
601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
2055 GATTTAGACATGCAGACTATCAAGATAACCAAGCTAAAAAATTAAGTGGTGGTCAAAA 2114  
621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
2115 AGAAAGCTGCTATAGGAATTCCTGTTCTGGGAACCCAAAGATACTGCTCTAGATGAA 2174  
641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrrAsnLeuLysTyrArg 660

2175 CCAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGTATGGAATCTTTAAAAATACAGA 2234  
661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
2235 AAAGCCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTTGA 2294  
681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
2295 GATAGGAAGCTGTGATATCACAGGAATGCTGAAATGTGTGGTTCTTCAATGTTCCTC 2354  
701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
2355 AAAAGTAATGGGGATCGGCTACGCCCTGAGCATGTATCATAGCAAAATTTTGTGCCACA 2414  
721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
2415 GAATCTCTTCTTCTACTGGTTAAACACATATACCTGGAGCTACTTTATTACACAGAA 2474  
741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
2475 GACCAACAACCTTGTGTATAGCTTGGCTTTCAAGGACATGGCAAAATTTTCAGGTTGTT 2534  
761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
2535 TCTGCCCTAGACAGCATTCATAATTTGGGTGGCATTTCTTATGGGGTTTCCATGAGACT 2594  
781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
2595 TTGGAAGAGCTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATAGT 2654  
801 ValPheThrGlnGlnProLeuGluMetAspSerLysSerPheAspGluMetGlu 820  
2655 GTATTTTACTCAGCAGCCACTGGAGGAGAAATGGATTCMAAATCTTTTGATGAATGGAA 2714  
821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
2715 CAGAGCTTACTTATCTTCTGAAACCAAGGCTKCTCTAGTGAGCACCATGAGCCCTTGG 2774  
841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
2775 AAACAACAGATGTATACATAGCAAAAGTTTCATTTCTTTTACCTTGAAAGCTGAAGTAA 2834  
861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
2835 TCAGTCAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTT 2894  
881 LeuValHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
2895 TTGGTTTCATCAGCTCTTTTAAAAATGCTGGTGGTCCCATCAAACTGTTCCAGACTTAT 2954  
901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
2955 TTTCTAAAACCTGGAGACAAACACATAAATACAAAACAAGCTGCTCTTCTCAAAATTC 3014  
921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
3015 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAAAGCCAGAACATATGTTGAC 3074  
941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
3075 ATGATTTAATGACAGTGAATGTATGTAACCGTGGCTCCCAATAGTGGGCTTTAAATGT 3134  
961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
3135 CATTCAGAAAAGGACATGTTTTTGCAGCTGTTTTTCAACAGTACTATGTTTATTCTTTA 3194  
981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
3195 CCTATATTAGTGAATATCATTAGTACTACTATCTTATCATTTAAATGTGACTGAACC 3254  
1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
3255 ATCCAGATCTGGAGTACCCCACTCTTTCAAGAAATTTACTGATATAGTTTTTAAAAATGAG 3314



```
QY 1021 LeuTyrPheGlnAlaLeuLeuGlyLeileValThrAlaMetProProTyrPheAla 1040
Db 3315 CTGTAATTTCAAGCAGCTTGGCTTGGAAATCATTTACTGCAATGCCACCTTACTTTGCC 3374
QY 1041 MetGluAsnAlaGluAsnHisLysIleValAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060
Db 3375 ATGGAAATCAGAGAAATCATAGATCAAGCTTAYACTCACTTAAACTTTCAGGTCCT 3434
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIle 1080
Db 3435 TTGCCATCTGCATATTGGATTGGACAAAGCTTGTGTGATATCCCTTATTTTATCAAT 3494
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100
Db 3495 CTTATTTTGTCTAGGAAGCTTATGGCAATTCATTTATGGATTATATTTTATCTGTA 3554
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120
Db 3555 AAGTTCTCTGCTGGTGTGCTTATGCTTATGTTATGTTATGTTCACT 3614
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140
Db 3615 TATATTGCTCTTTCACCTTTAAGAAAATTTTAAATACCAAGAAATTTTGGTCATTATC 3674
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160
Db 3675 TATTTCTGGCAGCGTTGGCTTGTATTGCAATCACTCAATAACTTTCATTATGGGATAC 3734
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180
Db 3735 ACAATTGCACTATTCTTCATTTATGCTTATGCTTATCAATCACTTCCAAATCTATCCACTCTA 3794
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200
Db 3795 GTTTCCTCCGATTTCTTTCATTAAGATTTCTTGGAGAAATGTACGAAATAATGTGGACACC 3854
QY 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220
Db 3855 TATAATCCATGGGATAGGCTTTCAGTAGCTGTTATATCGCTTACCTGCGAGTGTACTG 3914
QY 1221 TrpIlePheLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240
Db 3915 TGGATTTTCTCTTACAACTATGAGAAATAATATCGAGGCAGATCAATAAGAAAGAT 3974
QY 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260
Db 3975 CCTTTTTCAGAAACCTTTCAACGAAGCTTAAATAATAGGAAGCTTCCAGAACACCAGAC 4034
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280
Db 4035 AATGAGGATGAAGATGAAGATGTCAAAGCTGAAGACTAAAGGTCAAAGAGCTGATGGGT 4094
QY 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300
Db 4095 TGCCAGTGTGTGAGGAGAAACCATCCATTATGTCAGCAATTTGCATAAAGAATATGAT 4154
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320
Db 4155 GACAAGAAGAATTTTCTCTTTCAGAAAGAAAGTAAAGGTGCAACTAAATAATCATCTCT 4214
QY 1321 PheCysValLysLysGlyLeuIleLeuGlyLeuLeuGlyProAsnGlyValAlaGlyLysSer 1340
Db 4215 TTTCTGTGTGAATAAAGAGAGATCTTAGGACTATTGGTCCAAATGGTGTGGCAAGAAGC 4274
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360
Db 4275 ACAATTATTAATTTCTGTTGGTGATATTGAACCAACTTCAGGCCAGGATTTTATAGGA 4334
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380
Db 4335 GATTATTCTTCAGAGCAAGTGAAGATGATGATTCATGAAAGTGTATGGTTACTGTCTCT 4394
```

```
QY 1381 GlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyVala 1400
Db 4395 CAGATAAAACCCCTTTGTGTGCCAGATACTACATTCGACGAAACATTTTGAATTTATGGAGCT 4454
QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420
Db 4455 GTCAAGGAATGAGTGCAGTGAAGTGAAGTGCATTAAGTTCGAATTAACATGCACCT 4514
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440
Db 4515 GATTTTAAAGAACATCTTTCAAGAAGCTGTAAGAACAATACCTGCAGGAATCAACGAAGAAG 4574
QY 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460
Db 4575 TTGTGTTTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCACTCT 4634
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480
Db 4635 ACAGTATGGATCCCAAGGCCAAACAGCACATGTGGCGAGCAATTCGAACCTGCATTTAAA 4694
QY 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500
Db 4695 AACAGAAAGCGGCTGCTTATCTGCACCACTCACTATATGGAGGCGCAGAGGCTGTCTGT 4754
QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520
Db 4755 GATCGAGTAGCTATCATGCTGTCTGGCAGTTAAGATGTATCGGAACAGTACAACATCTA 4814
QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540
Db 4815 AAGATTAATTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAAC 4874
QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnIlePheProAsnAlaSerArgGln 1560
Db 4875 CTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCAAGCGCTCAG 4934
QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580
Db 4935 GAAAGTTTTTCTTATTTTGGCTTATTAATTTCTTAAGGAAGATGTTCACTCCCTTTCA 4994
QY 1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluTyrSerPhe 1600
Db 4995 CAATCTTTTTTAAGCTGGAAGAGCTAAACATGCTTTTGCATTGGAAGAAATATAGCTTT 5054
QY 1601 SerGlnAlaThrLeuGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620
Db 5055 TCTCAAGCAACATTCGAACAGGTTTTTGTAGAACTCACTAAAGAAACAAGAGAGGAAGAT 5114
QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640
Db 5115 AATAGTTGTGGAACTTTTAAACAGCACACTTTGTGGGAACGAAACAACAAGAGATAGAGTA 5174
QY 1641 ValPhe 1642
Db 5175 GTATTT 5180
RESULT 6
AB557750
ID AB557750 standard; cDNA; 4917 BP.
XX
AC AB557750;
XX
DT 04-FEB-2003 (first entry)
XX
DE Coding sequence of human ATP binding cassette ABCA5 transporter #1.
XX
KW Human; ATP binding cassette; ABC; ABCA5; transporter;
KW neurotoxin transport; beta-amyloid peptide; chromosome mapping;
KW blood brain barrier transport; tissue typing; predictive medicine;
KW ABCA5 mediated disorder; ABCA5 related disorder; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
```

FH Key Location/Qualifiers  
 FT 1. .4917  
 FT /\*tag= a  
 FT /product= "ABCAS"  
 FT /note= "ATP binding cassette (ABC) A5 transporter"  
 FT /transl\_except= (pos:3160..3162, aa:Ser)  
 XX US2002123107-A1.  
 XX  
 XX 05-SEP-2002.  
 XX  
 XX 01-MAR-2002; 2002US-00090458.  
 XX  
 XX 02-MAR-2001; 2001US-0272895P.  
 XX  
 XX (ACTI-) ACTIVE PASS PHARM INC.  
 XX  
 XX Chen H, Kilinski L, Le Bihan S;  
 XX WPI, 2003-066798/06.  
 XX P-FSDB; ABG72423.  
 XX  
 XX Novel isolated ATP binding cassette transporter family polypeptide,  
 XX ABCA5, useful for treating disorders associated with aberrant or unwanted  
 XX ABCA5 transporter expression or activity.  
 XX  
 XX Claim 2; Fig 1; 52pp; English.  
 XX  
 XX The invention describes an isolated ATP binding cassette (ABC)  
 XX transporter family polypeptide (I), designated ABCA5. (I) or the  
 XX polynucleotide encoding it (II) are useful as targets for developing  
 XX modulating agents to regulate a variety of cellular processes,  
 XX particularly the transport of neurotoxic molecules, e.g., beta-amyloid  
 XX peptide (Aβ), across cell membranes or, e.g., the blood brain barrier  
 XX (BBB), as targets for developing modulating agents of multi-drug  
 XX resistance, as diagnostic and therapeutic tools, or to treat disorders  
 XX associated with aberrant or unwanted ABCA5 transporter expression or  
 XX activity. (I), (II) or a host cell (III) expressing (II) are useful in  
 XX screening assays, detection assays (e.g., chromosomal mapping, tissue  
 XX typing, forensic biology), predictive medicine (e.g., diagnostic assays,  
 XX prognostic assays, monitoring clinical trials and pharmacogenomics), and  
 XX in methods of treatment (e.g., therapeutic and prophylactic). (I) or  
 XX (III) are useful as reagents or targets in assays applicable to treatment  
 XX and diagnosis of ABCA5-mediated or related disorders. (I) is useful to  
 XX screen for naturally occurring ABCA5 substrates; to screen for drugs or  
 XX compounds which modulate ABCA5 activity; as a bait protein in a yeast two  
 XX -hybrid or three-hybrid assay; and to identify other proteins which bind  
 XX to or interact with ABCA5. (II) is useful in: gene therapy; to detect  
 XX ABCA5 mRNA or a genetic alteration in a ABCA5 gene; to modulate ABCA5  
 XX activity; to locate gene regions associated with genetic disease or to  
 XX associate ABCA5 with the disease, to identify an individual from a minute  
 XX biological sample (tissue typing), and to aid in forensic identification  
 XX of the biological sample. This sequence encodes a novel human ATP binding  
 XX cassette (ABC) A5 transporter  
 XX  
 XX Sequence 4917 BP; 1540 A; 821 C; 955 G; 1599 T; 0 U; 2 Other;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 0 Length: 4917  
 XX Score: 8309.00 Matches: 1629  
 XX Percent Similarity: 98.85% Conservative: 1  
 XX Best Local Similarity: 98.79% Mismatches: 2  
 XX Query Match: 98.61% Indels: 18  
 XX DB: 7 Gaps: 2  
 XX  
 XX US-10-090-458-5 (1-1642) x ABS57750 (1-4917)  
 XX  
 XX QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
 XX  
 XX DB 1 ATGTCCACTGCATATAGGAGGTAGGAGTTTGGACAGACACCACTTCTACTGAAG 60  
 XX  
 XX QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGlnLeuLeuPheProLeu 40  
 XX  
 XX DB 1 ATGTCCACTGCATATAGGAGGTAGGAGTTTGGACAGACACCACTTCTACTGAAG 60

DB 61 AATTAATTAAATTAATGCAGAACCAAAAGAGTAGTAGTGTTCAGGAAATTCITTTTCCACTA 120  
 QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
 DB 121 TTTTITTTATTTTGGTTAAATTAATTAATAGCATGATGCATCCAAATAAGAAATATGAAGA 180  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
 DB 181 GTGCCTAATATAGAACTCAATCCTATGACAAAGTTTACTCTTTCTAATCTAATTTCTTGA 240  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 DB 241 TATACTCCAGTGACTAATATTACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA 300  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 DB 301 COTGATGTCATAATTACTGAAGAATATACAATGAAAGAAATGTTTAACATCCAGTCTC 360  
 QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
 DB 361 TCTAAGCCGAGCAACTTTTGTAGGTGTGTTTCAAGAGACTCCATGTCTCTATGAATTCGT 420  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 DB 421 TTTTITTCCTGATATGATTCACAGTATCTTCTATTATATGGATTCAAGAGCTGGCTGTTC 480  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerGlyPheThrValLeuGlnAlaSerIle 180  
 DB 481 AANTCATGTGAGGCTGCTCAGTACTGTGTCCTCAGGTTTTCACAGTATTATCAAGCATCC 540  
 QY 181 AspAlaAlaIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
 DB 541 GATGTGTCATATATACAGTTGAAGCAATGTTCTCTTTTGAAGAGAGCTGGAGTCAACT 600  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
 DB 601 AAGCTGTATTATGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGAGTA 660  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 DB 661 ATTTTAATATACCTAGTTATATAGCATTTTTCACCTTTTGGATCTTTTGGCAATTCATATC 720  
 QY 241 ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 DB 721 GTAGCGAAAAGAAAAGAAAATAAAGAAATTTTAAAGATAATGGACTTCATGATCT 780  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 DB 781 GCCTTTTGGCTTTCCTGGGTTCTCTATATACAGTATTAAATTTTCTTATGTCCCTCTT 840  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
 DB 841 ATGGCAGTCATTCGACAGCTTCTTTGTTATTCTCTCAAGTAGCAGCATTTGTGATATT 900  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 DB 901 CTGCTTTTTTCTCTTATGATTAATCATCTGATATTTTGTGTTTAAATGTGCACACTCTT 960  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 DB 961 TTTAAAAAATCAAAACATGTGGAAATAGTTGAATTTTGTACTGTGCTTTTGGATTT 1020  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 DB 1021 ATTGGCTTTATGATAATCTCTATAGAAAGTTTTCCAAATCGTTAGTGTGGCTTTTCAGT 1080  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 DB 1081 CCTTTCGTGACGTACTTTTGTGATGTTATGACAGTCTATGCATTTAGAGATTTT 1140  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 DB 1141 AATGAAGTGTCTTCATTTTCMAATTGTGACTGCAGGCCCATATCCTCTAATATTATTAAT 1200

401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 1201 ATCATGCTCACACTTAATAGTATATTCATGTCTCTTGCGTGTCTATCTTTGATCAAGTC 1260  
 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 1261 ATTCAGGGGAATTTGGCTTACGGAGATCATCTTATATTTTCTGAAGCCTTCATATTGG 1320  
 441 SerLysSerLysArgAsnTyrGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 1321 TCAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGCAATGTTAAATGGAATATTAGT 1380  
 461 PheSerGluIleLeuProValSerSerGluPheValGlyLeuAlaIleArgIle 480  
 1381 TTATAGTAAATATTAGCCAGCTTTCTTCAGAAATTTGAGGAAAGAACCCATGAAGATT 1440  
 481 SerGlyIleGlnLysThrTyrArgLysLysGlyLeuAsnValGluAlaLeuArgAsnLeu 500  
 1441 AGTGGTATTACAGAGACATACAGAAAGAGGGTGAATGTGGAGCCTTTGAGAAATTTG 1500  
 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 1501 TCATTTGACATATATGAGGGTCAGATTACTGCTTTACTTGGCCACAGTGGAAACAGGAAG 1560  
 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
 1561 AGTACATTGATGAATATCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGGATCTATA 1620  
 541 TyrGlyHisArgValSerGluLeuAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 1621 TATGACACACAGAGTCTCAAGAAATAGTAAATGTTTGAAGCAAGAAAAATGATTGGCAAT 1680  
 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 1681 TGTCACAGCTTAGATATACATCTTTGATGTTTGGACAGTAGAAGAAATTTATCAATTTTG 1740  
 581 AlaSerIleLysGlyIleProAlaAsnAsnIleLeuGlnGluValGlnLysValLeuLeu 600  
 1741 GCTTCAATCAAAGGATACACGCCAACAAATATAATACAAAGAGTGCAGAGGTTTTACTA 1800  
 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 1801 GATTAGACATCACAGCTATCAAGATACCAAGCTAAATAATTAAGTGGTGGTCAAAAA 1860  
 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
 1861 AGAAGCTGCTATTAGGAATTCCTGTTCTGGAAACCCCAAGATACTGCTGCTAGATGAA 1920  
 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
 1921 CCAACAGCTGGAATGGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAAAATACAGA 1980  
 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 1981 AAAGCCATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTTGCA 2040  
 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 2041 GATAGGAAGCTGTGATATCACAGGAATGCTGAAATGTTGGTTCCTTCAATGTTCTCTC 2100  
 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 2101 AAAAGTAAATGGGGATCGGCTACCGCTGACGATGTACATAGCAAAATATTGGGCCACA 2160  
 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
 2161 GAATCTCTTTCTTCACTGGTTAAACACACATATACCTGGAGCTACTTTATTACACAGAA 2220  
 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
 2221 GACCAACAACTTGTGTATAGTTCGCTTTCAGGACATGGCAAAATTTTCAGGTTTGTGTT 2280

761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
 2281 TCTGCCCTAGACAGCTATTCAAATTTGGGTGTCTATTCTTATGGTGTTCATGACGACT 2340  
 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
 2341 TTGGAAGACGATATTTTAAAGCTAGAGTTGAAGCAAGAAATGACCAAGCAGATTATAGT 2400  
 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
 2401 GTATTTACTCAGCAGCAGCTGGAGGAGAAATGGAATCAAATCTTTTATGATGAATGGAA 2460  
 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
 2461 CAGAGCTTACTTATCTTTCTGAACCAAGGCTGCTCTAGTGAGCAGCAGTGGCCCTTTGG 2520  
 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
 2521 AAACAACAGATGATACATAGCAAGTTTCACTTTCTTTACCTTGAACCTGGAAGTAA 2580  
 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTAAATTTTTTTCACAGTTCAGATTTTTATGTTT 2640  
 881 LeuValHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
 2641 TTGGTTCATCACTCTTTTAAATGCTGTGGTTCCTCATCAACTGCTTCCAGACTTATAT 2700  
 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
 2701 TTTCTAAACCTGGAGCAAAACACATAATAACAAACAAAGTCTGCTCTTCAAAATTTCT 2760  
 921 Ala-----AspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
 2761 GCTGGTGAGAGTGTNNGTGAAGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGC 2820  
 935 GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
 2821 CAGAAATATGTTGACGATGATTAATGACAGTGACTATGATTCCTGGTGCCTCCCATAGT 2880  
 955 AlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSer 974  
 2881 CGCGCTTTTAAATGTGATGATTCAGAAAAGGACTATGTTTTTGGACGTGTTTTCAACAGT 2940  
 975 ThrMetValTyrSerLeuProIleLeuValAsnIleLeuSerAsnTyrTyrLeuTyrHis 994  
 2941 ACTATGTTTATTTCTTACTATATTAGTGAATATCATTTAGTAACCTACTATCTTTATCAT 3000  
 995 LeuAsnValThrGluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAsp 1014  
 3001 TTAATGTGACTGAAACCATCCAGATCTCGAGTACCCTTCTTTCAAGAAATTTACTGAT 3060  
 1015 IleValPheLysIleGluLeuTyrPheGlnAlaAlaLeuLeuGly-IleIleValThrAl 1034  
 3061 ATAGTTTTTAAATTTGAGCTGTATTTTCAGCAGCTTTCCTTGGAAATCATTTGTAACCTGC 3120  
 1034 aMetProProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGln 1054  
 3121 AATGCCACCTTACTCTGCCATGGAAATGCAGAGATCATAA----- 3162  
 1054 nLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIle 1074  
 3163 -----GGTCTTTTGGCATCTGCATATTGGATTTGCAAGCTGTTGTTGATAT 3209  
 1074 eProLeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrGln 1094  
 3210 CCCCTTATTTTATCATTTCTTATTTGATGCTAGGAGCTTATTGGCATTTTCAATTATGG 3269  
 1094 YLeuTyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValPhe 1114  
 3270 ATTATATTTTATCTGTAAGTTCCTTCTGCTGGTTTTTTCCTTATTGGTATGTTCTCC 3299  
 1114 oSerValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLys 1134

```

Db      3330 ATCAGTATTCTGTTCACTTATATGCTTCTTTCACCTTTAAGAAAAATTTAAATACCAA 3389
Qy      1134 sGluPheTrpSerPheIleTySerValAlaAlaLeuAlaCysIleAlaIleThrGluIl 1154
Db      3390 AGAATTTTGGTCATTATCTATTCTGTGGCAGCGTTGGCTTGTATTGCAATCACTGAAT 3449
Qy      1154 eThrPhePheMetGlyTyThrIleAlaThrIleLeuHisTyValAlaPheCysIleIleIl 1174
Db      3450 AACTTTCTTTATGGGATACACAAATGCAACTATTCTTCATTATGCTTTTGTATCATCAT 3509
Qy      1174 eProIleTyProLeuLeuGlyCysLeuIleSerPheIleTySerTrpLysAsnVa 1194
Db      3510 TCCAATCTATCCACTTCTAGTGTGCTGATTTCTTTCATAAAGATTTCTTGGAGAATGT 3569
Qy      1194 lArgLysAsnValAspThrTyAsnProTrpAspArgLeuSerValAlaValIleSerPr 1214
Db      3570 ACGAAAAAATGGGACACCTATAATCCATGGGATAGGCTTTTCAGTAGCTGTATATCGCC 3629
Qy      1214 oTyTrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyTrpGluLysLysTyGlyGl 1234
Db      3630 TTACCTGCAGTGTGTAAGTCTGCTTCTTCTTACAACTACTATGAGAAAAAATATGGAGG 3689
Qy      1234 yArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLy 1294
Db      3690 CAGATCAATPAAGAAAGATCCCTTTTTCAGAAACCTTTCAACGAAAGTCTAAAAATAGGAA 3749
Qy      1254 sLeuProGluProProAspAsnGluAspGluAspValLysAlaGluAsgLeuLy 1274
Db      3750 GCTTCAGAACCCACCAGACAATGAGGATGAAGATGAGATGTCAAAGCTGAAGACTAAA 3809
Qy      1274 sValLysGluLeuMetGlyCysGlnCysCysGlnGluLysProSerIleMetValSerAs 1294
Db      3810 GGTCAAAGAGCTGATGGGTGTCAGTGTGTGAGGAGAAACCATCCATTATGGTCAGCAA 3869
Qy      1294 nLeuHisLysGluTyAspAspLysLysAspPheLeuLeuSerArgLysValLysLysVa 1314
Db      3870 TTTCATAAAGATATATGATGACAAAGAGATTTCTTTTTCAGAAAGATTAAGAAAGT 3929
Qy      1314 lAlaThrLysTrpIleSerPheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyPr 1334
Db      3930 GGCACATAAATACATCTCTTCTGTGTGAAAAAGAGAGAGATCTTAGGACTATTGGGTCC 3989
Qy      1334 oAsnGlyAlaGlyLysSerThrIleLeuAsnIleLeuValGlyAspIleGluProThrSe 1354
Db      3990 AATAGTGTCTGGCAAAAGCACAAATATTAAATATCTGGTGTGTGATATTGAACCAACTTC 4049
Qy      1354 rGlyGlnValPheLeuGlyAspTySerSerGluThrSerGluAspAspSerLeuLy 1374
Db      4050 AGCCACAGGTATTTTAGGAGATTTATTTCTTCAGAGACAAGTGAAGATGATGATTCAC 4109
Qy      1374 sCysMetGlyTyTyCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHi 1394
Db      4110 GTGTATGGGTACTGTCTCCATGATAAACCCTTTGTGGCAGATCTACTATGTCAGGAACA 4169
Qy      1394 sPheGluIleTyGlyValValLysGlyMetSerAlaSerAspMetLysGluValIleSe 1414
Db      4170 TTTTGAATTTATGGAGCTGTCAAGGATGAGTGCAGTGCATGAAGAAGTCATAAG 4229
Qy      1414 rArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuPr 1434
Db      4230 TCGAATTAACACATGCACTTGATTTAAAGAACATCTTCAGAGAGCTGTAAAGAACTACC 4289
Qy      1434 oAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleTh 1454
Db      4290 TGCAGGAATCAACGAAAGTGTGTGTTTCTCTAAGTATGCTAGGGAATCCTCAGATTAC 4349
Qy      1454 rLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAl 1474
Db      4350 TTTGTAGATGACCACTCTACAGGTATGGATCCCAAGCCAAACAGCACATGTGGCGAGC 4409
Qy      1474 aIleArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeuThrHisTyMetGl 1494

```

```

Db      4410 AATTCAACTGCATTTAAAAACAGAAAGCGGGCTGCTATTCTGCACCACCTACTATATGGA 4469
Qy      1494 uGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIl 1514
Db      4470 GGAGGCAGAGGCTGTCTGTGATCGAGTAGCTATCATGGTGTCTGGGCAGTTAAGATGTAT 4529
Qy      1514 eGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyTrpPheLeuGluIleLysLe 1534
Db      4530 CGAAACAGTACACATCTAAAGATTAATTTGGAAAAGSCTACTTTTGGAAATTAATTT 4589
Qy      1534 uLysAspTrpIleGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyTrpIlePh 1554
Db      4590 GAAGGACTGTGATAGAAAACCTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTT 4649
Qy      1554 eProAsnAlaSerArgGlnGlnSerPheSerSerIleLeuAlaTyTrpLysIleProLysGl 1574
Db      4650 CCCAATGACAGCGTCAGAAAGTTTTTCTTCTATTATTTGGCTTATAAATTCCTAAGGA 4709
Qy      1574 uAspValGlnSerLeuSerGlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAl 1594
Db      4710 AGATGTTCACTCCCTTTCACAAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGC 4769
Qy      1594 aIleGluGluTyTrpSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLy 1614
Db      4770 CATTGAAGATATAGCTTTTCTCAGCAACATTTGGAACAGGTTTTTGTAGAACTCACTAA 4829
Qy      1614 sGluGlnGluGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTrpGluAr 1634
Db      4830 AGAACAGAGGAGGAGAAATAGTTGTGGAACCTTTAAACAGCACACTTTGTGGGAACG 4889
Qy      1634 gThrGlnGluAspArgValValPhe 1642
Db      4890 AACACAAGAGATAGAGTAGTAGTTT 4914

RESULT 7
ABS57749
ID ABS57749 standard; cdna; 5463 BP.
XX
AC ABS57749;
XX
DT 04-FEB-2003 (first entry)
XX
DE cDNA encoding novel human ATP binding cassette ABCA5 transporter #1.
XX
KW Human; ATP binding cassette; ABC; ABCA5; transporter;
KW neurotoxin transport; beta-amyloid peptide; chromosome mapping;
KW blood brain barrier transport; tissue typing; predictive medicine;
KW ABCA5 mediated disorder; ABCA5 related disorder; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
CDS 114..5030
FT /*tag= a
FT /product= "ABCA5"
FT /note= "ATP binding cassette (ABC) A5 transporter"
FT /transl_except= (pos:3273..3275, aa:Ser)
XX
BN US2002123107-A1.
XX
PD 05-SEP-2002.
XX
PF 01-MAR-2002; 2002US-00090458.
XX
PR 02-MAR-2001; 2001US-0272885P.
XX
PA (ACTI-) ACTIVE PASS PHARM INC.
XX
FI Chen H, Kilinski L, Le Bihan S;
XX
DR WPI; 2003-066798/06.
DR P-PSDB; ABG72423.
XX

```

PT Novel isolated ATP binding cassette transporter family polypeptide,  
 PT ABCA5, useful for treating disorders associated with aberrant or unwanted  
 PT ABCA5 transporter expression or activity.

PS Claim 2; Fig 1; 52pp; English.

XX The invention describes an isolated ATP binding cassette (ABC)  
 CC transporter family polypeptide (I), designated ABCA5. (I) or the  
 CC polynucleotide encoding it (II) are useful as targets for developing  
 CC modulating agents to regulate a variety of cellular processes,  
 CC particularly the transport of neurotoxic molecules, e.g., beta-amyloid  
 CC peptide (Aβ), across cell membranes or, e.g., the blood brain barrier  
 CC (BBB), as targets for developing modulating agents of multi-drug  
 CC resistance, as diagnostic and therapeutic tools, or to treat disorders  
 CC associated with aberrant or unwanted ABCA5 transporter expression or  
 CC activity. (I), (II) or a host cell (III) expressing (II) are useful in  
 CC screening assays, detection assays (e.g., chromosomal mapping, tissue  
 CC typing, forensic biology), predictive medicine (e.g., diagnostic assays,  
 CC prognostic assays, monitoring clinical trials and pharmacogenomics), and  
 CC in methods of treatment (e.g., therapeutic and prophylactic). (I) or  
 CC (III) are useful as reagents or targets in assays applicable to treatment  
 CC and diagnosis of ABCA5-mediated or related disorders. (I) is useful to  
 CC screen for naturally occurring ABCA5 substrates; to screen for drugs or  
 CC compounds which modulate ABCA5 activity; as a bait protein in a yeast two  
 CC -hybrid or three-hybrid assay; and to identify other proteins which bind  
 CC to or interact with ABCA5. (II) is useful in: gene therapy; to detect  
 CC ABCA5 mRNA or a genetic alteration in a ABCA5 gene; to modulate ABCA5  
 CC activity; to locate gene regions associated with genetic disease or to  
 CC associate ABCA5 with the disease, to identify an individual from a minute  
 CC biological sample (tissue typing), and to aid in forensic identification  
 CC of the biological sample. This sequence encodes a novel human ATP binding  
 CC cassette (ABC) A5 transporter

XX SQ Sequence 5463 BP; 1699 A; 912 C; 1055 G; 1794 T; 0 U; 3 Other;

# Alignment Scores:

Pred. No.: 0 Length: 5463  
 Score: 8309.00 Matches: 1629  
 Percent Similarity: 98.85% Conservative: 1  
 Best Local Similarity: 98.79% Mismatches: 2  
 Query Match: 98.61% Indels: 18  
 Ds: 7 Gaps: 2

US-10-090-458-5 (1-1642) x ABS57749 (1-5463)

QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20  
 Db 114 ATGTCCTGCAATAGGAGGTAGGAGTTGGAGACAGACACCACTTCTACTGAAG 173  
 QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGlnIleLeuPheProLeu 40  
 Db 174 AATTACTTTAATAATGACAAACCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 233  
 QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTrpGluGlu 60  
 Db 234 TTTTITTTATTTGGTTAATATTAAATAGCATGATGCATCAATAGAAATATGAGAA 293  
 QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
 Db 294 GTGCGTAATATAGAACTCAATCCTATGGACAAGTTTACTCTTCTTAATCTAAATTTCTTGA 353  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 Db 354 TATACTCAGTGACTAATATTACAGCAGCATCATGAGAAAGTGTCTACTGATCATCTA 413  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGlnLysGluMetLeuThrSerSerLeu 120  
 Db 414 CCTGATGTCAATAATTACTGAAGATATACAAATGAAAGAAATGTTTAACATCCAGTCTC 473  
 QY 121 SerLysProSerAsnPheValGlyValPheLysAspSerMetSerTyrGluLeuArg 140  
 Db 474 TCTAAGCCGACCACTTTGTAGGTGTGTTTCAAGACTCCATGTCCTATGAACCTTCGT 533

QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
 Db 534 TTTTTCCTGATATGATTCAGTATCTCTTAATTTATGATTCAGAGCTGAGCTGCTTCA 593  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 Db 594 AATCATGTGAGGCTGCTCAGTACTGCTCAGGTTTCACAGTTTACAGCATCCATA 553  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
 Db 654 GATGCTGCCATATATACAGTTGAAGCAATGTTCTCTTTTGAAGAGAGCTGAGTCAACT 713  
 QY 201 LysAlaValIleMetGlyGluThrAlaValGluIleAspThrPheProArgGlyVal 220  
 Db 714 AAGCTGTTTATATGGAGAACTGCTGTGTAGAAATAGATACCTTTTCCCGAGAGTA 773  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 Db 774 ATTTTAATATACCTAGTTATAGCATTTTCCCTTTGGATCTTTTGGCAATTCAATC 833  
 QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
 Db 834 GTAGCAGAAAGAAAGAAATATAAGAAATTTTAAAGATATGGGACTTCATGATATC 893  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrTrpSerLeuIlePheLeuMetSerLeuLeu 280  
 Db 894 GCCTTTTGGCTTCTCTGGGTTCTCTATATACAGTTTAAATTTTCTTATGTCCTCTT 953  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
 Db 954 ATGGCAGTCATTTGGCAGCTTCTTTGTTTATTTCTCAAGTAGCAGCATTTGATATTT 1013  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerValPhePheAlaLeuMetLeuThrProLeu 320  
 Db 1014 CTGCTTTTCTTTTATGATATCACTGTATTTTGTCTTAAATGCTGACACTCTT 1073  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 Db 1074 TTTTAAATAATCAAAACATGCGGAATAGTTGAATTTTGTACTGTGCTTTTGGATTT 1133  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 Db 1134 ATTGGCTTATGATTAATCTCTATAGAAATTTTCCAAATCTGTAGTGGCTTTTCAGT 1193  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 Db 1194 CCTTCTGTCAGTACTGTTCTTGTGATTTGATTTGACAGGTCATGCAATTTAGAGATTT 1253  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 Db 1254 AATGAAGGTGCTTCAATTTTCAAAATTTGACTGAGGCCCATATCCTCTAATATTACAAT 1313  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 Db 1314 ATCATGCTCACACTTAATAGTATATTTCTATGTCCTTGGCTGTCTATCTTGTATCAAGTC 1373  
 QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTrp 440  
 Db 1374 ATTCCAGGGAAATTTGGCTTTAGGAGATCATCTTTATATTTTCTGAAGCTTTCATATTGG 1433  
 QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 Db 1434 TCAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGGCAATGTTAATGGAATATTAGT 1493  
 QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
 Db 1494 TTTAGTGAATTTATTTAGCCAGTTTCTTTCAGAAATTTGTAGGAAAGAGCCATAAGAAAT 1553  
 QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
 Db 1554 AGTGGTATTCAGAGACATACAGAAAGAGGGTGAATTTGGAGGCTTTGAGAAATTTG 1613  
 QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520

Db |||||TCATTTCACATATATGAGGGTCAGATTACTGCCCTTACTTGGCCACAGTGGAAACAGGAAG 1673  
Qy SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540  
Db 1674 AGTACATTGATGAATATCTTTGTGGACTCTGCCACCTCTGATGGTTTGCATCTATA 1733  
Qy TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db 1734 TATGGACACAGAGCTCAGAAATAGATCAAAATGTTTGAAGCAAGAAATGATGGCAAT 1793  
Qy CysProGlnLeuAspIleHisPheAspValLeuThrValGluGlnLeuSerIleLeu 580  
Db 1794 TGTCCACAGTTAGATATACCTTTGATGTTTGGACGTAGAGAAATTTATCAATTTG 1853  
Qy AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnValGlnLysValLeuLeu 600  
Db 1854 GCTTCAATCAAAAGGATACCAAGCAACAATATAACAAGAGTCAGAAAGTTTACTA 1913  
Qy AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys 620  
Db 1914 GATTAGACATGCAGACTATCAAGATACCAAGCTAAATAATTAAGTGGTGTCAAAA 1973  
Qy ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
Db 1974 AGAAAGCTGTCAATTAGGAATGCTGTTCTTGGGAACCCAAAGATACCTGCTAGATGAA 2033  
Qy ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyrArg 660  
Db 2034 CCAACAGCTGGAATGGACCTCTTCGACATATTTGTGGAATCTTTTAAATACAGA 2093  
Qy LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 2094 AAAGCCATCGGTGACAGCTTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTTGA 2153  
Qy AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2154 GATAGAAAGCTGTGATATCAAGAGATGCTGAATGTGTGTTCTTCAATGTTCCCTC 2213  
Qy LysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2214 AAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACA 2273  
Qy GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db 2274 GAATCTCTTCTTCACCTGGTTAAACACATATACCTGGAGCTACTTTATTACACAGAA 2333  
Qy AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 2334 GACCAACAACTGTGTATAGCTTGGCTTTCAAGACATGGCAAAATTTTCAGTTTGT 2393  
Qy SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2394 TCTGCCCTAGACAGCTCAATCAAAATTTGGTGTCTATTTCTATGGTGTTCATGACGACT 2453  
Qy LeuGluAspValPheLeuLysLeuGlnValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2454 TTGGAAAGACGTATTTTAAAGCTAGAGTTGAGCAGAAATTTGACCAAGCATATAGT 2513  
Qy ValPheThrGlnGlnProLeuGluGlnMetAspSerLysSerPheAspGluMetGlu 820  
Db 2514 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAATCTTTTGATGAAATGGAA 2573  
Qy GlnSerLeuLeuIleLeuSerGlnThrLysAlaSerLeuValSerMetSerLeuTrp 840  
Db 2574 CAGAGCTTACTTATTTCTTGTAAACCAAGCTGCTCTAGTGAGCACCATGAGCCTTTGG 2633  
Qy LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db 2634 AAACACAGATGTATACATAACAAAGTTTCAATTTCTTTTACCTTGAACGTTGAAGTAAA 2693  
Qy SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880

Db 2694 TCAGTGAGATCAGTGTCTGCTCTCTTTTAAATTTTTTTCACAGTTCAGATTTTATGTTT 2753  
Qy LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 2754 TTGGTTTCATCACTCTTTTAAAAATGCTGTGGTCCCATCAAACTGTTCAGACTTATAT 2813  
Qy PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
Db 2814 TTCTTAAACCTCGGAGACAAACCATTAATACAAACAACTCTGCTTCTTCAAAATCT 2873  
Qy Ala-----AspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
Db 2874 GCTGTGAGAGTGTGNGTGAAGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAGC 2933  
Qy GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
Db 2934 CAGAACATAAATGGTGACGATGATTAATCAGAGTACTATGATCGTGGCTCCCATAGT 2993  
Qy AlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSer 974  
Db 2994 GCGGCTTTAAATGTGATGCAATTCAGAAAAGGACTATGTTTTTTCAGCTGTTTTCAACAGT 3053  
Qy ThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHis 994  
Db 3054 ACTATGGTTTATCTTTACCTATATTTAGTGAATATCATTTAGTAACACTATCTTTATCAT 3113  
Qy LeuAsnValThrGluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAsp 1014  
Db 3114 TTAATGTGACTGAACCATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTTACTCAT 3173  
Qy IleValPheLysIleGluLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAl 1034  
Db 3174 ATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTAACG 3233  
Qy aMetProProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrG 1054  
Db 3234 AATGCACCTTACTCTGCCATGGAAATGCGAGAGATCATAA----- 3275  
Qy nLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspI 1074  
Db 3276 -----GGTCTTTTGGCATCTGCATTTTGGATTTGGACAAAGCTGTGTGTGATAT 3322  
Qy eProLeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrG 1094  
Db 3323 CCCCATTATTTTATCATCTTATTTTGTATGCTAGAGAGCTTATTTGGCATTTTCAATATG 3382  
Qy yLeuTyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValPr 1114  
Db 3383 ATTATATTTTATCTGTAAAGTTCCTTGTGCTGCTGTTTTTTTGGCTTATTTGGTTATGTC 3442  
Qy oSerValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLy 1134  
Db 3443 ATCAGTTATCTGTTCACCTTATTTGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCA 3502  
Qy sGluPheTrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluI 1154  
Db 3503 AGAATTTTGTTCATTTATCTATTTCTGTGGCAGCGTGTGGCTTGTATTTGCAATCAGTGA 3562  
Qy eThrPhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIle 1174  
Db 3563 AACTTTCTTTATGGGATACAAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCAT 3622  
Qy eProIleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnVa 1194  
Db 3623 TCCAACTATCCACTTCTAGTTGCTGATTTCTTTTCAATAAGATTTCTTGGAAAGATGT 3682  
Qy lArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSerPr 1214  
Db 3683 ACGAAAAATGGACACCATATATTCATGGGATAGGCTTTTTCAGTACTGTTTATATCGCC 3742  
Qy oTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysTyrGlyG 1234  
Db 3743 TTACTTCGAGTGTGACTGTGGATTTTCTTCACTACTATGAGAAAAATATGGAGG 3802







Db 1921 CCAACAGCTGGAATGGACCCCTGTTTCGACATATTTGATGGAATCTTTTAAAAATACAGA 1980  
Qy 661 LysAlaIleAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 1981 AAACCAATCCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATCTTGCA 2040  
Qy 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2041 CATAGGAAGCTGTGATATCAACAGGAATGCTGAAATGTGTGGTCTTCAATGTTCTC 2100  
Qy 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2101 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACA 2160  
Qy 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db 2161 GAACTCTCTTCTCACTGGTTAAACACATATATCTGGAGCTACTTTATTACACAGAA 2220  
Qy 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 2221 GACCAACAACTTGTGTATAGCTTGCCTTCAAGGACATGGACAAATTTTCAGGTTTGT 2280  
Qy 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2281 TCTGCCCTAGACAGCTCATTCAAATTTGGGTGGCAATTTCTTATGGGTTCATGACACT 2340  
Qy 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2341 TTGGAAGACGCTATTTTAAAGCTAGAAGTTGAAGCAAAATTTGACCAAGCAGATATAGT 2400  
Qy 801 ValPheThrGlnGlnProLeuGluGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2401 GTATTTACTAGCAGCCACTGAGAGGAAGAAATGGATTCAAAATCTTTTGATGAATGGAA 2460  
Qy 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 2461 CAGAGCTTACTTATTTCTTCTGAAACCAAGCTKCTCTAGTAGCAGCACCATGAGCCTTGG 2520  
Qy 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
Db 2521 AAACAACAGATGATACAATAGCAAGATTTTCATTTCTTTACCTTGAAACGCTGAAGATAA 2580  
Qy 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2581 TCAGTGAGATCAGTGTTGCTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTT 2640  
Qy 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 2641 TTGGTTTCATCACTCTTTTAAAAATGCTGTGGTTCCCATCAAACTGTGTCCAGACTTATAT 2700  
Qy 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 2701 TTTCTAAAACCTGGAGACAAACCAACATATAACAAACAGTCTGCTTTCTTCAAAATCT 2760  
Qy 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 2761 GCTGACTCAGATATCAGTGTATCTTATTAGCTTTTTCAAAGCCAGAACATATAATGGTGACG 2820  
Qy 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 2821 ATGATTAATGACAGTACTATGATATCCGTGGCTCCCATAGTGGGCTTTAAATGTGRTG 2880  
Qy 961 HisSerGluLysAspTyrValPheAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 2881 CATTCAGAAAAGGACTATGTTTTTGCAGCTGTTTTCAACAGTACTATGTTTATTTCTTTA 2940  
Qy 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 2941 CCTATATTAGGAATATCATTAGTACTATCTTATCTTATCATTTAATATGACTGAAC 3000  
Qy 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020

Db 3001 ATCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTTACTGATATAGTTTAAAAATTCAG 3060  
Qy 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 3061 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTACTGCAATGCCACTTTACTTTGCC 3120  
Qy 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3121 ATGGAAAAATGCAGAGATCATATAAGATCAAGACTTAYACTCAACTTAAACTTTTCAGTCTT 3180  
Qy 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
Db 3181 TTGCCATCTGCATATTTGGATTGGACAGCTGTGTGTATATCCCTTATTTTATCATTT 3240  
Qy 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3241 CTTATTTTGTGCTAGGAAGCTTATTTGGCAATTTTCATATGGATATATATTTTATCTGA 3300  
Qy 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 3301 AAGTTTCTTGTGTGGTTTTTGGCTTATTTGTTATGTATGTTCCATCAGTTATTTCTGTTCACT 3360  
Qy 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 3361 TATATTTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAATTTTGTTCATTTATC 3420  
Qy 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3421 TATTTCTGTGGCAGCGCTTGGCTTGTATGCAATCACTGAAATAAATCTTTTATGGATAC 3480  
Qy 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 3481 ACAATTTGCACTATTTCTTTCATTTATGCTTTTGTATCATCTTCATCTATCCACTTCTA 3540  
Qy 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
Db 3541 GGTTCCTCTGATTTCTTTCATAAAGATTTCTTGAAGAATGTACGAAAAAATGTGACACC 3600  
Qy 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 3601 TATAATCCATGGATAGGCTTTTCAGTAGCTGTATATTCGCTTTACCTCAGTGTGACTG 3660  
Qy 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 3661 TGGATTTTCTCTTACAATCTATGAGAAAAAATATGAGGAGCAGATCAATAAGAAAAAGAT 3720  
Qy 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 3721 CCTTTTTCAGAAAACTTTCAACGAAGCTTAAAAATAGGAAGCTTCCAGAACCCACAGAC 3780  
Qy 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 3781 AATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTTAAAGGTCAAAGAGCTGATGGGT 3840  
Qy 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 3841 TGCCAGTGTGTGAGAGAAACCACTCCATATGTGTGTCAGCAATTTGCAATAAAGATATAT 3900  
Qy 1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320  
Db 3901 GACAAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 3960  
Qy 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValAlaGlyLysSer 1340  
Db 3961 TTCTGTGTGAAGAAAGGAGAGATCTTAGGACTATTTGGGTCCAAATGGTGGCGGCAAAAGC 4020  
Qy 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 4021 ACAATTTATATTTCTGGTTGGTGATATTCGAACCACTTCAGGCCAGGTATTTTAGGA 4080  
Qy 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 4081 GATTATTTCTCAGACACAAGTCAAGATGATGATTTCACTGAAGTGTATGGGTACTGTCTCT 4140

QY 1381 GlnIleAspProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
 DB 4141 CAGATAAACCCCTTGTGGCCAGATACATCATTTGCAGGAACATTTTGAATAATTTATGGAGCT 4200  
 QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
 DB 4201 GTCAAGAGGAAATGAGTCAAGTACATGAAGAAGTCATAAGTCGAATACACATGCACATT 4260  
 QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
 DB 4261 GATTTAAAGAAACATCTTCAGAGAAGCTGTAAAGAAACTACCTGCAGGAATCAAAACGAAG 4320  
 QY 1441 LeuCysePheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
 DB 4321 TTGTGTTTTCTCTAAGTATGTAGGAATCTTCAGATTACTTTCTAGATGAACCATCT 4380  
 QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
 DB 4381 ACAGGTATGGATCCCAAGCCAAACAGACACATGTGGCGAGCAATTCGAACCTGCATTAA 4440  
 QY 1481 AsnArgLysArgAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500  
 DB 4441 ACAGAAAGCGGCTGCTATTCTGACCACTCATATATGGAGAGCGAGCGCTGCTGT 4500  
 QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
 DB 4501 GATCGAGTAGCTATCATGTGCTGGCGAGTTAAGATGTATCGGAACAGTACACATCTA 4560  
 QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
 DB 4561 AAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAATTAAGAGACTGGATAGAAAC 4620  
 QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
 DB 4621 CTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCAAGCGCTCAG 4680  
 QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
 DB 4681 GAAAGTTTTCTCTTATTTGGCTATATAAATTCCTAAGGAAGATGTTTCAGTCCCTTTCA 4740  
 QY 1581 GlnSerPhePheLysLeuGluGlu 1588  
 DB 4741 CAAATCTTTTTTAAGCTGGAGAA 4764

## RESULT 9

ADA53416

ID ADA53416 standard; cDNA; 3268 BP.

XX AC ADA53416;

XX DT 20-NOV-2003 (first entry)

XX DE Human coding sequence, SEQ ID 984.

XX XX

XX KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;

XX KW Gene Therapy; human; secretory protein; membrane proteins; cancer;

XX KW Inflammatory disease; osteoporosis; neurological disease; gene; ss.

XX OS Homo sapiens.

XX XX

XX PN EP1293569-A2.

XX XX

XX PD 19-MAR-2003.

XX XX

XX PF 21-MAR-2002; 2002EP-00006586.

XX XX

XX XX

XX PR 14-SEP-2001; 2001JP-00328381.

XX PR 24-JAN-2002; 2002US-0350435P.

XX XX

XX PA (HELI-) HELIX RES INST.

XX PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX XX

PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;  
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;  
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;  
 XX WPI: 2003-395539/38.  
 DR P-PSDB; ADA55055.  
 XX PT New polynucleotides encoding full-length polypeptides, e.g. secretory  
 PT and/or membrane proteins, useful for developing medicines for diseases in  
 PT which the gene is involved, or as target molecules for gene therapy.  
 XX PS Claim 1; SEQ ID NO 984; 205pp; English.  
 XX CC The present invention relates to novel human secretory or membrane  
 CC proteins (ADA54072-ADA55110) and their coding sequences (ADA52433-  
 CC ADA54071). The coding sequences are useful in the gene therapy of  
 CC diseases caused by abnormalities of the proteins, e.g. cancer,  
 CC inflammatory diseases, osteoporosis or neurological disease.  
 XX SQ Sequence 3268 BP; 1022 A; 540 C; 622 G; 1084 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0 Length: 3268  
 Score: 4847.50 Matches: 941  
 Percent Similarity: 98.64% Conservative: 2  
 Best Local Similarity: 98.43% Mismatches: 2  
 Query Match: 57.53% Indels: 11  
 DB: 7 Gaps: 1

US-10-090-458-5 (1-1642) x ADA53416 (1-3268)

QY 698 MetPheLeuLysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyr 717  
 DB 6 ATGTTCTCTCAAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGCAAAATAT 65  
 QY 718 CysAlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeu 737  
 DB 66 TGTGCCACAGAAATCTCTTCTTCACCTGGTTAAACAATATACCTGGAGCTACTTTATTA 125  
 QY 738 GlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSer 757  
 DB 126 CAACAGAAATGACCACAACTTGTGTATAGCTTGCCTTCAAGGACATGACAAATTTTCA 185  
 QY 758 Gly-----LeuPheSerAlaLeuAspSerHis 766  
 DB 186 GGAATTCGTTGTATAGACAAAGGATTTATGATGTTTCTTCCTCCCTAGACATCAT 245  
 QY 767 SerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeu 786  
 DB 246 TCANATTTGGTGTCTATTTCTTATGTTTCCATGACGACTTGGAGACGTTATTTT 305  
 QY 787 LysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnPro 806  
 DB 306 AAGCTAGAAGTTGAAGCAGAAATTTGACCAAGCAGATTTAGTGTATTTTACTCAGACCCA 365  
 QY 807 LeuGluGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuLeu 826  
 DB 366 CTGGAGAGAAATGATTCAAATCTTTTGTGAAATGGAACAGAGCTTACTTATCTT 425  
 QY 827 SerGluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGlnMetTyrThr 846  
 DB 426 TCTGAAACCAAGGCTTCTCTAGTGAGCACCAGCTTTGGAAACAACAGATGTATACA 485  
 QY 847 IleAlaLysPheHisPhePheThrLeuLysArgGluSerLysSerValArgSerValLeu 866  
 DB 486 ATAGCAAAAGTTTCTATTTCTTACCTTGAAACGTGAAGTAATTCAGTGAGATCAGTGTG 545  
 QY 867 LeuLeuLeuLeuPhePheThrValGlnIlePheMetPheLeuValHisSerPhe 886  
 DB 546 CTTCTGCTTTTAAATTTTTCACAGTTCAAGATTTTATGTTTGTGTTTCTCATCTCTTT 605  
 QY 887 LysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLysProGlyAsp 906

Db 606 AAAATGCTGGTCCCATCAAACTGTTCCAGACTTATATTTCTAAACCTGGAGAC 665  
Qy 907 LysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAspIleSer 926  
Db 666 AAACACATAAATACAAACCAAGCTGCTCTCTCAAAATTCCTGCTGACTCAGATATCAGT 725  
Qy 927 AspieuIleSerPhePheThrSerGlnAsnIleMetValThrMetIleAsnAspSerAsp 946  
Db 726 GATCTTATTAGCTTTTTCACAGGCAGAACATAATGGGACGATGATTAATGACAGTGAC 785  
Qy 947 TyrValSerValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyr 966  
Db 786 TATGTATCCGTGGCTCCCATAGTGGCGCTTTAAATGTGATGCATTCAGAAAAGCACTAT 845  
Qy 967 ValPheAlaAlaValPheAsnSerThrMetValTyrSerLeuPheProIleLeuValAsnIle 986  
Db 846 GTTTTTCAGCTGTTTTCACACAGTACTATGTTTATTTTAAATTCAGCTGATTTTCAAGCAGTC 905  
Qy 987 IleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTyrSerThr 1006  
Db 906 ATTAGTAACTACTATCTTTATCATTTAATGTGACTGAACCATCCAGATCTGGAGTACC 965  
Qy 1007 ProphePheGlnGluIleThrAspIleValPheIysIleGluLeuTyrPheGlnAlaAla 1026  
Db 966 CCATCTCTTCAAGAAATTAAGTATAGTATTTTAAATTCAGCTGATTTTCAAGCAGCT 1025  
Qy 1027 LeuLeuGlyIleIleValThrAlaMetProTyrPheAlaMetGluAsnAlaGluAsn 1046  
Db 1026 TTGCTTGAATCATTTGCTGCAATGCCACTTACTTTGCCATGGAAATTCGACAGAAAT 1085  
Qy 1047 HisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrp 1066  
Db 1086 CATAAGATCAAAAGCTTATACCTCAACTTAAACTTTTCAGGCTCTTTGCCATCTGCATATTTGG 1145  
Qy 1067 IleGlyGlnAlaValAlaAspIleProLeuPheIleIleLeuIleLeuMetLeuGly 1086  
Db 1146 ATTGACCAAGCTGGTGTGATATCCCTTATTTTATCATCTTATTTTATGATGCTAGGA 1205  
Qy 1087 SerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaVal 1106  
Db 1206 AGCTTACTGGCATTTTCATTTATGATATATTTTATATCTGTAAGATTCCTTGCTGGTT 1265  
Qy 1107 PheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPheThr 1126  
Db 1266 TTTTGCCCTTATGGTTATGTTCCATCAGTATTCGTTCATCTATATTTGCTTCCTTCCAC 1325  
Qy 1127 PheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyrSerValAlaAlaLeu 1146  
Db 1326 TTTAAGAAATTTTAAATACCAAGAAATTTTGGTCACTTATCTATCTTCTGTCGACGCTTG 1385  
Qy 1147 AlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyrThrIleAlaThrIleLeu 1166  
Db 1386 GCTTGTATTGCAATCACTGAATAACTTTCTTTATGGGATACACAAATTCGAACATTTCTT 1445  
Qy 1167 HisTyrAlaPheCysIleIleProIleProIleTyrProLeuLeuGlyCysLeuIleSerPhe 1186  
Db 1446 CATTTAGCCTTTTGTATCATCTTCATCTATCCACTTCACTTCAAGTGTGCTGATTTCTTTC 1505  
Qy 1187 IleLysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnProTrpAspArg 1206  
Db 1506 ATAAAGATTTCTTGAAGAATCTACGAAAAATGTGGACACCTATAATCCATGCCGATAGG 1565  
Qy 1207 LeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGln 1226  
Db 1566 CTTTCAGTAGCTTATATTCGGCTCACCTCGAGTGTGACTGTGAGTTTCTCTTACAA 1625  
Qy 1227 TyrTyrGluLysLysTyrGlyGlyArgSerIleArgLysAspProPhePheArgAsnLeu 1246  
Db 1626 TACTATGAGAAATAATGAGGACGACATCAATAAGAAAGATCCCTTTTTCAGAAACCTT 1685  
Qy 1247 SerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAspGluAspGlu 1266  
Db 1686 TCAACCAAGCTCTAAAAATAGGAAGCTTCCAGAACCCAGACCAATGAGGATGAAGATGAA 1745

Qy 1267 AspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGlu 1286  
Db 1746 GATGTCAAAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGTTCAGGTGTGTGAGGAG 1805  
Qy 1287 LysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeu 1306  
Db 1806 AAACCATTCATTTATGGTTCAGCAATTTGCATAAAGAAATATGATGACAGAAAGATTTCTT 1865  
Qy 1307 LeuSerArgLysValLysValLysValAlaThrLysTyrIleSerPheCysValLysLysGly 1326  
Db 1866 CTTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTCTGTGTGAAAAAGGA 1925  
Qy 1327 GluIleLeuGluLeuLeuGlyProAsnGlyAlaGlyLysSerThrIleIleAsnIleLeu 1346  
Db 1926 GAGATCTTTAGGACTATTTGGGTCCAAATGGTGTGTCGCAAAAGCACAAATTTATTTATCTGT 1985  
Qy 1347 ValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThr 1366  
Db 1986 GTTGGTGATATTGAACCAACTTCAGGCCAGGTATTTTAGGAGATTTATTTCTTCAGAGACA 2045  
Qy 1367 SerGluAspAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTrp 1386  
Db 2046 AGTGAAGATGATGTTTCACTGAAGTGTATGGGTACTGTCTCAGATAAACCCCTTTGTGG 2105  
Qy 1387 ProAspThrThrLeuGlnGluHisPheGluIleTyrGlyValaValLysGlyMetSerAla 1406  
Db 2106 CCAGATCTACTATTCAGAGNACATTTTGAATTTATGGAGCTGTCAAGAGATGAGTGCA 2165  
Qy 1407 SerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeu 1426  
Db 2166 AGTGACATGAAAGAAAGTCAATAGTGCATTAACACACTGGACTTGATTTAAAGAACATCTT 2225  
Qy 1427 GlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSer 1446  
Db 2226 CAGAAAGCTGTAAAGAAACTTACTCTGAGGAATCAACAGAAAGTGTGTGTTTCTCTAAGT 2285  
Qy 1447 MetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLys 1466  
Db 2286 ATGCTAGGGAATCCTCAGATTTACTTTGTAGATGAACCATCTACAGGTATGATCCCAA 2345  
Qy 1467 AlaLysGluHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaAla 1486  
Db 2346 GCCAAACAGCACATGTGGCGAGCAATTCGACTGCATTTAAATAACAGAAAGCGGCTGT 2405  
Qy 1487 IleLeuThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMet 1506  
Db 2406 ATTCTGACCACCTCACTATATGAGGAGGAGGAGGCTGTCTGTGATCGAGTAGCTATCATG 2465  
Qy 1507 ValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLys 1526  
Db 2466 GTGCTGGGCGAGTTAAGATGTATCGGAACAGTACACATCTTAAGAGTAGATTTGGAAAA 2525  
Qy 1527 GlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeuGluValAspArgLeu 1546  
Db 2526 GGCTACTTTTGGAAATTTAAATTGAAGGACTGATAGAAAAACCTAGAAAGTAGACGCCCTT 2585  
Qy 1547 GlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSerPheSerSerIle 1566  
Db 2586 CAAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGCTCAGGAAAGTTTTTCTTCTATT 2645  
Qy 1567 LeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeu 1586  
Db 2646 TTGGCTTATAAATTCCTTAAGGAAGATGTTAGTCCCTTTTCAACTCTTTTAAAGCTG 2705  
Qy 1587 GluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThrLeuGlu 1606  
Db 2706 GAAGAGAGCTAAACATGCTTTTGCATTGAAGAATATAGCTTTTCTCAGGCAACATTTGGAA 2765  
Qy 1607 GlnValPheValGluLeuThrLysGluGlnGluGluAspAsnSerCysGlyThrLeu 1626  
Db 2766 CAGTTTTTGTAGAACTCTCTAAAGAACCAAGAGGAGAGATAATAGTTGTGGAACCTTTA 2825

QY 1627 AsnSerThrLeuTrpGluArgThrGlnGluAspArgValValPhe 1642  
 Db 2826 AACAGACACTTGGTGGACGACACACAGAGATAGAGTAGTATT 2873

## RESULT 10

ADCS1606

ID ADCS1606 standard; DNA; 2723 BP.

XX

XX ADCS1606;

XX 18-DEC-2003 (first entry)

XX

DE Human macroprotein-94.64 coding sequence, SEQ ID 1.

XX

XX Human; macroprotein-94.64; dementia; arrhythmia; asthma; diabetes; gene;

KW ds.

XX

XX Homo sapiens.

XX

XX Key

FT CDS

FT Location/Qualifiers

FT 12..2486

FT /\*tag= a

FT /product= "Human macroprotein-94.64"

XX

XX CN1382726-A.

XX

XX 04-DEC-2002.

XX

XX 26-APR-2001; 2001CN-00112765.

XX

XX 26-APR-2001; 2001CN-00112765.

XX

XX (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.

XX

XX Mao Y, Xie Y;

XX

XX WPI; 2003-269487/27.

XX

XX P-PSDB; ADCS1607.

XX

XX New human macroprotein-94.64, encoding polynucleotide, antagonist and

PT recombinant production, useful for treating dementia, arrhythmia, asthma

PT and diabetes.

XX

XX Claim 6; SEQ ID NO 1; 35pp; Chinese.

XX

XX The present invention relates to human macroprotein-94.64 (ADCS1607) and

CC its coding sequence (ADCS1606). The protein is useful for treating

CC dementia, arrhythmia, asthma and diabetes.

XX

XX Sequence 2723 BP; 855 A; 460 C; 518 G; 890 T; 0 U; 0 Other;

XX

XX Alignment Scores:

Pred. No.: 0

Score: 4282.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 50.82%

DB: 9

Length: 2723

Matches: 827

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-10-090-458-5 (1-1642) x ADCS1606 (1-2723)

QY

816 PheAspGluMetGluGlnSerLeuLeuLeuSerGluThrLysAlaSerLeuValSer 835

Db

3 TTGTGTAAGTAAGACAGAGCTTACTTATCTTCTGAACCAAGGCTTCTAGTGAGC 62

QY

836 ThrMetSerLeuTrpLysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeu 855

Db

63 ACCATGAGCTTTGGAAACACAGATGTATACAAAGAGTTTCATTTCTTACCTTG 122

QY

856 LysArgGluSerLysSerValArgSerValLeuLeuLeuLeuLeuLeuLeuLeuVal 875

Db

123 AAACGTGAAAGTAAGTACAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTT 182

QY 876 GlnIlePheMetPheLeuValHisHisSerPheLysAsnAlaValValProIleLysLeu 895  
 Db 183 CAGATTTTATGTTTGGTTCATCACTCTTTTAAAAATGCTGGTGGTCCCATCAACTT 242  
 QY 896 ValProAspLeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeu 915  
 Db 243 GTTCCAGACTTATATTTCTAAACCTGGAGACAAACCAATAAAATACAAAACAAGCTG 302  
 QY 916 LeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGln 935  
 Db 303 CTTCCTCAAAATTCGCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAACCAG 362  
 QY 936 AsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSerAla 955  
 Db 363 AACATATGTCAGCATGATTAATGACAGTACTATGATCCGTGGCTCCCATAGTGGC 422  
 QY 956 AlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThr 975  
 Db 423 GCTTTAAATGTGATTCAGAAAGGACTATGTTTTCAGAGTGTTCACAGTACT 482  
 QY 976 MetValTyrSerLeuProIleLeuValAsnIleLeuSerAsnTyrTyrLeuTyrHisLeu 995  
 Db 483 ATGGTTTATCTTACCTATATTAGTGAATATCAATTAGTAACCTACTATCTTATCATTTA 542  
 QY 996 AsnValThrGluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIle 1015  
 Db 543 AATGTGACTGAACCATCCAGATCTGGAGTACCCCATCTTTTCAGAGAAATTACTGATATA 602  
 QY 1016 ValPheLysIleGluLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMet 1035  
 Db 603 GTTTTAAAAATGAGCTGTATTTCAGCAGCTTTCGTTGGAATCATTTGTTACTGCAATG 662  
 QY 1036 ProProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeu 1055  
 Db 663 CCACCTTACTTTGCCATGGAAATGAGAGATCATAGATCAAGCTTATCTACTCAACT 722  
 QY 1056 LysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIlePro 1075  
 Db 723 AAACCTTTCAGGCTTTTTCGCTCATCTGCATATTGGATTGGACAGCTGTTGTTGATATCCC 782  
 QY 1076 LeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeu 1095  
 Db 783 TTATTTTATCATCTTATTTTGTGCTAGGAGCTTACTGGCATTTTTCATTATGATTA 842  
 QY 1096 TyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSer 1115  
 Db 843 TATTTTATCATCTGTAAGTTCCTTGTGCTGTTTTCGCTTATGTTATGTTCCATCA 902  
 QY 1116 ValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGlu 1135  
 Db 903 GTTATTCTGTTCATTTATTTGCTTTTCCCTTTAAGAAAATTTTAAATACCAAGAA 962  
 QY 1136 PheTrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThr 1155  
 Db 963 TTTTGTGCTATTTATCTTCTGTCGAGGCTTGGCTTGTATTGCAATCATCAATAACT 1022  
 QY 1156 PhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIlePro 1175  
 Db 1023 TTCTTTATGGGATACACAATTCACATTTCTTCATTTATGCTTTTGTATCATCTATCCA 1082  
 QY 1176 IleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArg 1195  
 Db 1083 ATCTATCCACTTCTAGTTGCTGATTTCTTTCATNAAGATTTCTTGGAGATGTACGA 1142  
 QY 1196 LysAsnValAspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyr 1215  
 Db 1143 AAAAATGTGACACCTTATATCCATGGGATAGGCTTTCAGTACTGTTTATCGCCTTAC 1202  
 QY 1216 LeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArg 1235  
 Db 1203 CTCAGTGTGCTGTGGATTTTCTTCTTACATCTATGAGAAAAAATATGGAGGACGA 1262  
 QY 1236 SerIleArgLysAspProphePheArgAsnLeuSerThrLysSerLysAsnArgLysLeu 1255

1263 TCATTAAGAAAGATCCCTTTTTCAGAAACCTTTTCAAGAGCTTAAATAAGAGAGCTT 1322  
 1256 ProGluProProAspAsnGluAspGluAspValLysValAlaGluLeuLysVal 1275  
 1323 CCAGAACCAACACAGACAAATGAGGATGAAGATGAAGATGTCAGAGCTGAAGAGTCA 1382  
 1276 LysGluLeuMetGlyCysGlnCysCysGluGluLysProSerIleMetValSerAsnLeu 1295  
 1383 AAGAGCTGATGGTGGTCCAGTGGTGGTGGGAGAAACCATCCATATGTCAGCAATTTG 1442  
 1296 HisLysGluTyrAspAspLysLysPhePheLeuLeuSerArgLysValLysValAla 1315  
 1443 CATAGAATATGATGACAAAGAAATTTCTTTTCAAGAAAGTAAAGAGTGGCA 1502  
 1316 ThrLysTyrIleSerPheCysValLysLysGlyGluLysLeuGlyLeuGlyProAsn 1335  
 1503 ACTAAATACATCTCTTTCTGTGTGAAAGAGAGATCTTAGGACTATTGGTCCAAAT 1562  
 1336 GlyAlaGlySerThrIleLeuValGlyAspIleGluProThrSerGly 1355  
 1563 GGTGCTGGCAAGACAAATTAATATTCTGTTGGTGTATTTGAACCACTTCAGGC 1622  
 1356 GlnValPheLeuGlyAspTyrSerGluThrSerGluAspAspSerLeuLysCys 1375  
 1623 CAGGTATTTTAGGAGATTTATCTTCAGAGACAAGTGAAGATGATTCATCTGAAGTGT 1682  
 1376 MetGlyTyrCysProGlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPhe 1395  
 1683 ATGGGTACTGTCTCTCAGATTAACCTTTGTGGCCAGATCTACATTCAGGAAATTTT 1742  
 1396 GluIleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArg 1415  
 1743 GAAATTTATGGAGCTCTCAAGGAATCAGTGAAGTGCATCAAGAAAGTCAATAGTGA 1802  
 1416 IleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAla 1435  
 1803 ATAACATGTCATCTGATTTAAAGAACATCTTCAGAGACATGTTAAAGAACTACTGCA 1862  
 1436 GlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeu 1455  
 1863 GGAATCAACGAAGTTGGTGTCTTCTAAGTATGTAGGGAATCCTCAGATTAATCTTG 1922  
 1456 LeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIle 1475  
 1923 CTAGATGAACCACTCTCAGTATGATGATCCCAAGCCAAACAGCAGCATGTGGGAGCAAT 1982  
 1476 ArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGlu 1495  
 1983 CGAACTGCATTTTAAACAGAAAGCGGCTGTATTTCTGACCACCTCATATATGGAGGAG 2042  
 1496 AlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGly 1515  
 2043 GCAGAGGCTGTCTGTATCGAGTATCATGTATCATGTCTGGGAGTAAAGATGTATCGGA 2102  
 1516 ThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLys 1535  
 2103 ACAGTACACATCTTAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTAAG 2162  
 1536 AspTrpIleGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePhePro 1555  
 2163 GACTGGATGAAACCTTAGAAGTAGACCGCTTCAAGAGAGAAATTCAGTATATTTTCCCA 2222  
 1556 AsnAlaSerArgGlnGluSerPheSerIleLeuAlaTyrLysIleProLysGluAsp 1575  
 2223 AATGCAAGCCGTGAGGAAGTTTCTTCTATTTTGGCTTATAAATTCCTAAGAGAT 2282  
 1576 ValGlnSerLeuSerGlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIle 1595  
 2283 GTTCAGTCCCTTTCAACAATCTTTTAAAGTGAAGAGCTAAACATGCTTTTGCAT 2342  
 1596 GluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGlu 1615

2343 GAAGATATAGCTTTTCTCAAGCAACATTTGGACAGGTTTTTGTAGAACTCACTAAAGAA 2402  
 1616 GlnGluGluGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTyrTrpGluArgThr 1635  
 2403 CAAGAGGAGGAAGATATAGTTGTGAACCTTTAAACAGACACATTTTGGTGGGAACGAACA 2462  
 1636 GlnGluAspArgValValPhe 1642  
 2463 CAAGAGATAGAGTAGTATTTT 2483

RESULT 11  
 ADA20300  
 ID ADA20300 standard; cDNA; 5797 BP.  
 XX  
 AC ADA20300;  
 20-NOV-2003 (first entry)  
 XX Human ATP-binding cassette transporter A8 (ABCA8) cDNA sequence.  
 DE Human; ATP-binding cassette transporter A8 protein; ABCA8 protein;  
 XX active transporter protein; ion transport; amino acid transport;  
 KW peptide transport; sugar transport; vitamin transport;  
 KW steroid hormone transport; antilipemic; ophthalmological;  
 KW neuroprotective; cardiac; gene therapy; hypercholesterolemia;  
 KW drug resistance; retinal degeneration transport; neurological disease;  
 KW elevated cholesterol level; neurological disease;  
 KW retinal degeneration transport; drug resistance; heart disease;  
 KW chemotherapy; ss; gene; chromosome 17q24.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 139..5004  
 FT /\*tag= a  
 FT /product= "Human ABCA8 protein"  
 FT 1927..2046  
 FT /\*tag= b  
 FT /note= "This sequence of this region was newly identified  
 in the specification and specifically claimed in claim 9"  
 XX  
 XX misc\_feature  
 XX WO2003064591-A2.  
 XX  
 XX 07-AUG-2003.  
 XX  
 XX 18-OCT-2002; 2002WO-US033496.  
 XX  
 XX 19-OCT-2001; 2001US-0351004P.  
 XX  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX (AVET ) AVENTIS PHARMA.  
 XX  
 XX Dean MC, Schriml L, Allikmets RL, Deneffe MP, Rosier-Montus M;  
 PI Arnould-Reguigne I, Prades C;  
 XX P-PSDB; ADA20301.  
 XX  
 XX WPI; 2003-618356/58.  
 XX  
 XX New ATP-binding cassette transporter A8 (ABCA8) proteins and nucleic acid  
 PT molecules, useful for diagnosing or treating diseases associated with  
 PT altered expression of ABCA8, e.g. hypercholesterolemia or neurological  
 PT disease.  
 XX  
 XX Claim 9; Fig 4; 114pp; English.  
 XX  
 XX This invention relates to a novel substantially purified (functional)  
 CC human ATP-binding cassette transporter A8 (ABCA8) protein. The protein,  
 CC encoded by a gene located on chromosome 17q24, is an active transporter  
 CC protein associated with transport of various substrates, for example  
 CC ions, amino acids, peptides, sugars, vitamins or steroid hormones.  
 CC Compounds which modulate the activity of the protein of the invention may  
 CC have antilipemic, ophthalmological, neuroprotective or cardiac  
 CC activities. The DNA sequence which encodes the protein of the invention









Db 684 TGGAGATAAAAGAG--AGTTTATGGATGGACAAATAACATGTTCTAGCGCTATAT 741  
Qy 46 LeuileuLeuSerMetMetHisProAsnLysLysTyrgluValProAsnleGlu 65  
Db 742 TTGTGCATCTTTCCGGAACACATTCAGAGCTACCGCTTTTCTCGAACACCTCTAAAGTC 801  
Qy 66 LeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGlyTyrrProValThr 85  
Db 802 CTGGGAAGCGTGATCAGTTAATAGACTCTGGCTGGTAGTGCCATATACACCACTCAGT 861  
Qy 86 AsnileThrSerSerileMetGlnLysValSerThrAspHisLeu-----ProAsp 102  
Db 862 AACATAACCAAGGATAATGAATAAGATGGCGTTCCTTTATGAAGAGAACAA 921  
Qy 103 ValileleThrGluGluTyrrAsnGluLysGluMetLeuThrSerSerLeuSerLys 122  
Db 922 GTCATTTGGACACCATGAT-----GAAGACCATGGATATAGAACTTCCCAAAAAA 972  
Qy 123 ProSerAsnPheValGlyValPheLysAspSerMetSerTyrgluLeuArgPhe--- 141  
Db 973 TACCATGAATGGTGGAGTATATTTAGTGATCTTTCTCATATCGCTGAAGTTAAT 1032  
Qy 142 PheProAspMetileProValSerSerileTyrrMetAspSerArgAlaGlyCys----- 159  
Db 1033 TGGGGATATAGATCCAGTTATAAGGAGCACTCTGAATACACAGAACACACTGTGGGCC 1092  
Qy 160 -----SerLysSerCysGluAlaAlaGlnTyrrTrpSerSerGlyPheThrValLeu 176  
Db 1093 ATGCATGGTGMAATTTTGTGTACTTGGCAAGTACTGGCTAAAGGGTTTGTAGCTTTT 1152  
Qy 177 GlnAlaSerileAspAlaAlaileleGlnLeuLysThrAsnValSerLeuTyrrPheGlu 196  
Db 1153 CAAGCTGCAATTAATCGCAATTTAGAGATCACAACAAATCATCTGTATGGAGGAG 1212  
Qy 197 LeuGluSerThrLysAlaValileMetGlyGluThrAlaValValGluileAspThrPhe 216  
Db 1213 TTGACATCAGTATTGGATAAATATG-----AAGATACCACCTTTC 1254  
Qy 217 ---ProArgGlyValileLeu-----IleTyrrLeuValileAlaPheSer 230  
Db 1255 ATTTCTAAGGGAGAAATTAAGATGAATGGTTTTCATTTTACTTGTCTAGTTCTTCTCT 1314  
Qy 231 PropheGlyTyrrPheLeuAlaileHisileValAlaGluLysLysLysLysGlu 250  
Db 1315 TCTTTTATATCTTTCATCATTAAT---GTTGCAAGGAAAGAGGAAATTTAAGAA 1371  
Qy 251 PheLeuLysileMetGlyLeuHisAspThrAlaPheTrpLeuSerTrpValLeuLeuTyrr 270  
Db 1372 CTGACACAGTAATGGGTCTCCGAGAGTCAGCATTTCTGGCTCTCTGGNGATTCACATAC 1431  
Qy 271 ThrSerLeuilePheLeuMetSerLeuLeuMetAlaValileAlaThrAlaSerLeuLeu 290  
Db 1432 ATTTGGCTTCATCTTCAATATGTCATTTTATGGCTCTGGCTAATCAATCAATCTCAAT 1491  
Qy 291 PheProGlnSerSerSerileValilePheLeuLeuPhePheLeuTyrrGlyLeuSerSer 310  
Db 1492 GTATTTTCATCTGGCTTCATGTGTATTCACACTCTATAGCTTATATGGCTTCTTTG 1551  
Qy 311 ValPhePheAlaLeuMetLeuThrProLeuPheLysLysSerLysHisValGlyileVal 330  
Db 1552 ATAGCATTTGGCTTCTCTCATGAGTGTTTTAATAGGAAACCTATGCTCGCTGGTTGGCT 1611  
Qy 331 GluPhePheValThrValAlaPheGlyPheileGlyLeuMetileleLeuileGluSer 350  
Db 1612 GGATTTCTCTTCACTGTATTTGGGGATGCTCGGGA---TTCACTGTGTATACAGACAA 1668  
Qy 351 PheProLysSerLeuValTrp-----LeuPheSerProPheCysHisCysThrPhe 367  
Db 1669 CTTCCTTTATCTTTGGGATGGGTATTAAGTCTTCTTAGCCCTTTT-----GCCTTC 1719  
Qy 368 ValileGlyileAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSerPheSer 387

Db 1720 ACTGCTGGATGGCCAGGTTACACACTGGATTAATTAATTAAGTGTGTATTTTCTCT 1779  
Qy 388 AsnLeuThrAlaGlyProTyrrProLeuileleThrileleMetLeuThrLeuAsnSer 407  
Db 1780 GATCCCTCTGGGATTCATACAAAATGATAGCCACTTTTTCATTTTGGCATTTGATACT 1839  
Qy 408 IlePheTyrrValLeuLeuAlaValTyrrLeuAspGlnValileProGlyGluPheGlyLeu 427  
Db 1840 CTTTTCATTTGTATTCATATATTTGGCGAGTTTTACCTGTATGAAGTGGCCAT 1899  
Qy 428 ArgArgSerLeuTyrrPheLeuLysProSerTyrrTrpSerLysSerLysAsnTyrr 447  
Db 1900 GGGGATTTCTCATTTTCTTAAGTCTCAATTTTGGTCCAAACATCAAAATCACTCAT 1959  
Qy 448 GluGluLeuSerGluGlyAsnValAsnGlyAsnLysPheSerPheSerGluileleGluPro 467  
Db 1960 CATGAATCTTTGAGAAATGAATAAATCTCGAGCATTTCTCTGATGATTTCTTTGAACCG 2019  
Qy 468 ValSerSerGluPheValGlyLysGluAlaileArgLysSerGlyileGlnLysThrTyrr 487  
Db 2020 GTGTCTCCAGATTTCCATGGAAAGAGCCATTAAGATCAGAAATGTTATAAAGATAT 2079  
Qy 488 ArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAspLysTyrrGluGly 507  
Db 2080 AATGGAAGAGCTGAAAGTAGAAGCATTCGAAGGCATATTTTTCATATATGAAGGA 2139  
Qy 508 GlnileThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnileLeu 527  
Db 2140 CAGATCACTCAATCTTGGGCATTAATGAGCTGTAATCAACACTGCTTAACATCTT 2199  
Qy 528 CysGlyLeuCysProProSerAspGlyPheAlaSerileTyrrGlyHisArgValSerGlu 547  
Db 2200 AGTGGATTTGTCTTTCTACAGAGGATCAGCCACTATTTATTAATTAATCACTCTCTGAA 2259  
Qy 548 IleAspGluMetPheGluAlaArgLysMetileGlyileCysProGlnLeuAspLysHis 567  
Db 2260 ATAACGTGACATGGAAGAAATTAAGAAATATTTGGATTTGTCCAGATTCATTTTCAA 2319  
Qy 568 PheAspValLeuThrValGluGluAsnLeuSerileleAlaSerileLysGlyilePro 587  
Db 2320 TTTGACTTCTCTCTGTGAGAGAAACCTCAGCGGTATTTGCTAAAAATAAAGGGATTCAG 2379  
Qy 588 AlaAsnAsnileleGlnGluValGlnLysValLeuLeuAspLeuAspMetGlnThrile 607  
Db 2380 CCAAGGAAGTGAACACAGAGGTAAAGAAATTAATGAATTAAGATTAAGATTCGAAGCAT 2439  
Qy 608 LysAspAsnGlnAlaLysLysLeuSerGlyGlnLysArgLysLysLeuSerLeuGlyile 627  
Db 2440 CAAGACATTTATCTTAAATAATTAAGTGGTGGCAGAGAGAACTAAACACTAGGAT 2499  
Qy 628 AlaValileGlyAsnProLysLysileleLeuLeuAspGluProThrAlaGlyMetAspPro 647  
Db 2500 GCCATCTTAGAGATCTCAGGTTTGTGCTAGATGAACCACTCTCGATTTGGATCCC 2559  
Qy 648 CysSerArgHisileValTrpAsnLeuLeuLysTyrrArgLysAlaAsnArgValThrVal 667  
Db 2560 TTTTCAAGACACCGAGTGTGGAGCTCTCTGAAGGACATAAAGTAGACCGACTTATCCTC 2619  
Qy 668 PheSerThrHisPheMetAspGluAlaAspLysLeuAlaAspArgLysAlaValileSer 687  
Db 2620 TTCAGTACCAATTCATGATGAGGTGATCTTGGCTGTATGAGGAATATTTCTGTCT 2679  
Qy 688 GlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLysTrpGlyileGly 707  
Db 2680 AATGGGAAGTGAATATGTCAGGATCATCTTGTCTTCTGAAGGAAGTGGGTATTTGGA 2739  
Qy 708 TyrrArgLeuSerMetTyrrileAspLysTyrrCysAlaThrGluSerLeuSerLeuVal 727  
Db 2740 TATCATTTAAGTTTACACAGGAATGAATGTGTGACACAGAAATAATCACTCCCTTAT 2799  
Qy 728 LysGlnHisileleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnLeuValTyrrSer 747  
Db 2800 AACACGACATTTCTGATGCCAAGTTAAACACAGAAAGTGAAGAAAACTTTGTATATAGT 2859

QY 748 LeuProPheLeuAspMetAspLysPheSerGlyLeuPheSerAlaLeuAspSerHisSer 767  
DDB 2860 TTGGCTTTGGAAAAACGAAACAAATTTCCAGATCTTTACAGTGACCTTGATAAGTGTCT 2919  
QY 768 AsnLeuGlyValIleSerThrGlyValSerMetThrThrLeuGluAspValPheLeuLys 787  
DDB 2920 GACCAGGCAATAGGAATATCTGTTTCAGTGACATCTCTGAATGAAGTATCTTGAC 2979  
QY 788 LeuGluValGluAlaGluLeuAspGlnAlaAspTyrSerValPheThrGlnGlnProLeu 807  
DDB 2980 CTAGAAGGAAATTCAGCAATTCATGAACCAAGATTTTGACATGGGAAACAAAGAAATA 3039  
QY 808 GluGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuLeuSer 827  
DDB 3040 CATGTGACAGAAATACTCGAGATGAGTCTGAATAGGAACAGCTTTTGTTCTCTCT 3099  
QY 828 GluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGlnMetThrThrIle 847  
DDB 3100 GAAACAAAGAAAGCT--GTGAGTAGTCAGCTCTCTGGAGACGACAAATCTATGACGTG 3156  
QY 848 AlaLysPheHisPhePheThrLeuLysArgGluSerLysSerValArgSerValLeuLeu 867  
DDB 3157 GCAACACTTCGCTCTTAAAGTTAAGCGCTGAAGAGAGAGCTCTTTGTGTGTGTACTA 3216  
QY 868 LeuLeuLeuLeuPhePheThrValGlnIlePheMetPheLeuValHisHisSerPheLys 887  
DDB 3217 GTACTTGGATTCGCTTTT--ATCCCATCATCTTAGAGAAAGATATGATATAAGTAACT 3273  
QY 888 AsnAlaValProIleLysLeuValProAspLeuTyrPheLeuLysProGlyAspLys 907  
DDB 3274 CGTGAAACTCAATTTGGGAGTTTCCACCAGATGATATTTCTCTCTCGGAAACAAATC 3333  
QY 908 ProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAspIleSerAsp 927  
DDB 3334 CCGAAGACGCTCTTACCGCTGTTTAACTGTTATATACAGGATCAATATTCAGAC 3393  
QY 928 LeuIleSerPhePheThrSerGlnAsnIleMetValThrMet-----Ile 942  
DDB 3394 CTGCTGATTCACCTGAAGTGTGAGGATATAGTTTGGAAATAGATGACTTTAGAACACAGA 3453  
QY 943 AsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMetHisSer 962  
DDB 3454 AATGGCTCAGATGATCCCTCC-----TACATGGAGCCATCATAGTGTCTGGTAC 3504  
QY 963 GluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeuProIle 982  
DDB 3505 CAGAGGATTAACAGATTTCTGTTGCTGTGTAATACCAAGAAATGGAATTTCTCTGT 3564  
QY 983 LeuValAsnIleLeuSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIleGln 1002  
DDB 3565 CTTATGGGAATTTGATGCAATGCCCTTATGGGAATTTTAACTTCACGAGCTTATTCAA 3624  
QY 1003 IleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGluLeuTyr 1022  
DDB 3625 ACGGAGAGCACTTCATTTCTCTGTAT-----GACATAGTGTGATCTTGTTTATA 3678  
QY 1023 PheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAlaMetGlu 1042  
DDB 3679 GATGGGTCCATATTTTGTGTTGATCACAACACTGGCTTCTCTCTTTATTCGCGCATGAC 3738  
QY 1043 AsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuPro 1062  
DDB 3739 AGCATCAGCATATATAAATAATGTTCAATCCAGTATGATGATTCAGCCCTCTGGCT 3798  
QY 1063 SerAlaTyrTrpIleGlyGlnAlaValAspIleProLeuPhePheIleLeuIle 1082  
DDB 3799 TCAGTACTGTTGTGACAGGCTCTGTTGGACATTCATATTAATCTTCTGATTT----- 3852  
QY 1083 LeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPhe 1102  
DDB 3853 -----CTCTTTTCAATACATTTAATTTACTTCTATATTTCTGGGATTC 3897

QY 1103 -----LeuAlaValValPheCysLeuIleGlyTyrValProSer 1115  
DDB 3898 CAGCTTTTCATGGGAACATGTTTGGTGGTATGATTAATGTTTGTGCGAGTTCT 3957  
QY 1116 ValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGlu 1135  
DDB 3958 CTTATATCTCTCACAATGCTGCTTTTCAATCATCTTTCGCAAGTGGAGAAAAATAATGCG 4017  
QY 1136 PheTrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThr 1155  
DDB 4018 TTTTGGCTTTTGGCTTTTATTTATTT-----ATCTTAATATGTTATCCACAATATGGTATCA 4074  
QY 1156 PhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIle-----Ile 1174  
DDB 4075 -----ACTCAATATGAAAAACTCACTTAATTTTGTGTCATGATTTTCATA 4119  
QY 1175 ProIleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnVal 1194  
DDB 4120 CTTCTCTCTCTTGTGCGGTATGTCATGTTATTTATTCAGCTCGACTTT----- 4170  
QY 1195 ArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSer----- 1208  
DDB 4171 ATGAGAACTTGGACAGCTCTGGAC-----AATAGATAAATGAAGTCAATAAACCACT 4224  
QY 1209 ---ValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyr 1227  
DDB 4225 CTTTAAACAACCTTAATACCATACCTTCAGAGTGTATTTTCTCTTTTGTGCATAAGGTGT 4284  
QY 1228 TyGluLysTyrGlyArgSerIleArgLysAspProPhePheArgAsnLeuSer 1247  
DDB 4285 CTGGAATAGAGTATGGAATGAATGAATGAATGAATGAATGAATGAATGAATGAATGAAT 4341  
QY 1248 ThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAspGluAspGluAsp 1267  
DDB 4342 CCACGGAGTAGAGAACTCAT---CCCAATCCGGAAGACCCGGAAGAGATGAAGAT 4398  
QY 1268 ValLysGluGluLeuLysValLysGluLeuMetGlyCysGlnCysGluGluLys 1287  
DDB 4399 GTTCAAGCTGAAAGAGTCCAGAGCAAGCAATGCACTCACTGCTCCAACTTGGAGAGGAA 4458  
QY 1288 ProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeu 1307  
DDB 4459 CCAGTCATACTGCAAGCTGTTTACACAGGAATATATGACACAAAGAAAGTGTCTTT 4518  
QY 1308 SerArgLysValLysValAlaThrLysTyrIleSerPheCysValLysGlyGlu 1327  
DDB 4519 TCAACAAGAAAGAGAAATAGCCATCAGAAATGTTTCTCTTTTGTGTTAAAAAGGTGAA 4578  
QY 1328 IleLeuGlyLeuLeuGlyProAsnGlyValaGlyLysSerThrIleIleAsnIleLeuVal 1347  
DDB 4579 GTTTTGGATTTACTAGACACATGGAGCTGGTAAAAAGTACTTCCATTAATAATGATAACT 4638  
QY 1348 GlyAspIleGluProThrSerGlyGlnValPheLeu---GlyAspTyrSerSerGluThr 1366  
DDB 4639 GGGTGCACAAAGCAACTCGAGGAGTGGTGTTCACAGGCGACAGAGCATCAGTAAGG 4698  
QY 1367 SerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTyr 1386  
DDB 4699 CAACAGCATGACACACCTCAAGTCTTGGGTACTGCTCCCTCAGGAGAACTCACTCTGG 4758  
QY 1387 ProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAla 1406  
DDB 4759 CCCAAGCTTACATGAAAGAGCACTTGGAGTGTGATGTCAGCTGTGAAAGAGCTGCGCAA 4818  
QY 1407 SerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeu 1426  
DDB 4819 GAAGATGCTGCTCAGTATTTTACGATTTGGTGAAGCTCTTAAGCTCCAGGAACAATCT 4878  
QY 1427 GlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSer 1446  
DDB 4879 AAGCTCTGCTGTAACAACTCTATCAGAGGAATGAAGAGAAAGCTGCTTTGTGCTGAGC 4938  
QY 1447 MetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLys 1466

```

Db 4939 ATCTGGGGAACCATGAGTGTGCTTCTAGATGAGCGCTTACCGGATGACCCCGAG 4998
QY 1467 AlAlysGlnHieMetTrrArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaAla 1486
Db 4999 GGGCAGCAGCAAAATGGCAGATACCTCAGCGTACCGTTAAATAACAAGGAGGGGCACC 5058
QY 1487 IleLeuThrThrHieTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMet 1506
Db 5059 CTTCTGACCAACCATTAATGTCAGAGGCTGAGGCTGTGTGACCGGTATGGCCATGATG 5118
QY 1507 ValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLys 1526
Db 5119 GTGTGAGGAACGCTAAGGTGTATTGGTTCCATTCACATCTGAAACAAAGTTGGTAGA 5178
QY 1527 GlyTyrPheLeuGluIleLysLeuLysAspTrrIleGluAsnLeuGluValAspArgLeu 1546
Db 5179 GATTATTACTAGATAATAAATAATGAAA-----GAACCTACCAGGTGGAAGCTCTC 5229
QY 1547 GlnArgGluIleGlnTrrIlePheProAsnAlaSerArgGlnGluSerPheSerSerIle 1566
Db 5230 CACACAGAGATTTCAGAGCTTTTCCACAGCGTGTCTGCGAGGAAGATATCTCTCTTA 5289
QY 1567 LeuAlaTrrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeu 1586
Db 5290 ATGGCGTATAGTTACCTGTGAGGATGTCCACCCCTCTATCTCGGGCCTTTTCAAGTTA 5349
QY 1587 GluGluAlaLysHisAlaPheAlaIleGluGluTrrSerPheSerGlnAlaThrLeuGlu 1606
Db 5350 GAGGCGATGAAACACACCTTCACCTGGAGGATACAGCCCTCTCTCAGGCTACCTTGGAG 5409
QY 1607 GlnValPheValGluLeuThrLysGluGlnGluGluAspAsnSerCysGlyThrLeu 1626
Db 5410 CAGGTATCTTAGAACTCTGTAAGAGCAG--GAGCTGGGAATGTTGATGATAAAAT 5466
QY 1627 AsnSerThrLeuTrrTrrGlu---ArgThrGlnGluAsp 1638
Db 5467 GATACAAAGTTGAATGGAACTTCTCCACAGGAAGAC 5505

RESULT 13
AB222923
ID AB222923 standard; cDNA; 5722 BP.
XX AC AB222923;
XX AC AB222923;
XX DT 08-APR-2003 (first entry)
XX DE Human ABCA10 transporter encoding cDNA SEQ ID NO:3.
XX KW Human; ATP binding cassette transporter family; ABCA10 transporter;
XX KW chromosome 17; chromosome 17q24; gene; ss.
XX OS Homo sapiens.
XX FH Key
XX CDS 254..4885
XX FT /tag= a
XX FT /product= "ABCA10 transporter"
XX PN WO2002103015-A2.
XX PD 27-DEC-2002.
XX PF 13-JUN-2002; 2002WO-CA000873.
XX PR 14-JUN-2001; 2001US-0298604P.
XX PA (ACTI-) ACTIVE PASS PHARM INC.
XX PI Le Bihan S, Chen H, Killinski L;
XX DR WPI; 2003-167521/16.
DR P-FSDB; ABF56906.

```

XX New ABCA10 transporter nucleic acid molecules and polypeptides, useful in  
 PT screening assays or in predictive medicine, e.g. diagnostic assays,  
 XX prognostic assays, monitoring clinical trials or pharmacogenetics.  
 XX  
 PS Claim 1; Page 106-113; 117pp; English.

CC The present sequence encodes a human ATP binding cassette (ABC)  
 CC transporter, designated ABCA10 transporter (1). (1) is located on human  
 CC chromosome 17, more specifically to locus 17q24. (1) can be used in gene  
 CC therapy. The ABCA10 transporter nucleic acid molecules, proteins and  
 CC antibodies can be used in screening assays or in predictive medicine,  
 CC e.g. diagnostic assays, prognostic assays, monitoring clinical trials or  
 CC pharmacogenetics. The ABCA10 transporter nucleic acid molecules and  
 CC polypeptides are useful as targets for developing modulating agents that  
 CC regulate cellular processes, particularly the transport of neurotoxic  
 CC molecules across cell membranes. The nucleic acid molecules and  
 CC polypeptides are useful in treating a subject having a disorder  
 CC associated with aberrant or unwanted ABCA10 transporter protein or  
 CC nucleic acid expression or activity

XX SQ Sequence 5722 BP; 1806 A; 987 C; 1153 G; 1776 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 4.1e-252 Length: 5722  
 Score: 3229.00 Matches: 713  
 Percent Similarity: 61.36% Conservative: 313  
 Best Local Similarity: 42.64% Mismatches: 561  
 Query Match: 38.32% Indels: 86  
 DB: 7 Gaps: 30

US-10-090-458-5 (1-1642) x AB222923 (1-5722)

QY 7 GluValGlyValTrrArgGlnThrArgThrLeuLeuLeuLysAsnTrrLeuIleLysCys 26  
 Db 1 GAGATAAGCGTGGTCAACAAATTCAGGCTCTCTGTACAGAAATTTCTTAAATAATGS 60  
 QY 27 ArgThrLysLysSerValGlnGluIleLeuPheProLeuPhePheLeuPheTrrLeu 46  
 Db 61 AGAATAAAAGAG--AGTTTATTGGATGACAAATAACATTTCTTCTAGGGCTATATTG 118  
 QY 47 IleLeuIleSerMetMetHisProAsnLysLysTrrGluValProAsnIleGluLeu 66  
 Db 119 TGCATCTTTTGGGAACACTTCAGAGTACCGCTTCTTCTGAAACACTCTCTAAAGCTTG 178  
 QY 67 AsnProMetAspLysPheThrLeuSerAsnLeuLeuLeuGlyTrrThrProValThrAsn 86  
 Db 179 GGAAGCGTGGATCAGTTTAATGACTCTGGCTGTGTAGTGGCATATACACAGTCAGTAAC 238  
 QY 87 IleThrSerSerIleMetGlnLysValSerThrAspHisLeu-----ProAspVal 103  
 Db 239 ATAACACAAAGGATATGATAGATGGCTTGGCTTCTTTATGAAAGGAAGACAGTC 298  
 QY 104 IleIleThrGluLufyTrrAsnGluLysGluMetLeuThrSerSerLeuSerLysPro 123  
 Db 299 ATTGGACACCCAGAT-----GAAGAGACCATGGATATAGAACTTCCAAAAAATAC 349  
 QY 124 SerAsnPheValGlyValValPheLysAspSerMetSerTrrGluLeuArgPhe---Phe 142  
 Db 350 CATGAAATGGTGGAGTTATATTAGTACTTCTCATATCGCTGAAGTTTAATTGG 409  
 QY 143 ProAspMetIleProValSerSerIleTrrMetAspSerArgAlaGlyCys----- 159  
 Db 410 GGATATAGAAATCCAGTTATAAAGGAGCACTCTGTAATACACAGAACACTGTGGGCCATG 469  
 QY 160 -----SerLysSerCysGluAlaAlaGlnTrrTrrSerSerGlyPheThrValLeuGln 177  
 Db 470 CATGGTGAATTTTTTGTACTTGGCAAGTACTGGCTAAAGGGGTTTGTAGCTTTTCAA 529  
 QY 178 AlaSerIleAspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrrLysGluLeu 197  
 Db 530 GCTGCAATTAATGCTCAATTATAGAAGTCACACAAATCATCTCTGTAATGGAGGAGTTG 589

Qy	198	GlusSerThrLysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPhe---	216
Db	590	ACGTCAGCTTATGGAAATAAATATG-----AAGATACACCACTTTCATT	631
Qy	217	ProArgGlyValIleLeu-----IleTyrLeuValIleAlaPheSerPro	231
Db	632	TCTAAGGGAGAAATTATGAATGAATGCTTCACTTTTACCTTAGTTCTTCTCTCTCT	691
Qy	232	PheGlyTyrPheLeuAlaIleHisIleValalaGluLysGluLysIleLysGluPhe	251
Db	692	TTTATATACCTTTGCATCATTAAT--GTTGCAAGGAGAAAGAGAAATTTAAGAAACTG	748
Qy	252	LeuLysIleMetGlyLeuHisAspThrAlaPheTrpLeuSerTrpValLeuLeuTyrThr	271
Db	749	ATGACAGATAGGGTCTCCGAGAGTACAGATCTGGCTCTCCCTGGGATTCACATACAT	808
Qy	272	SerLeuIlePheLeuMetSerLeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPhe	291
Db	809	TGCTTCATCTTCATATGTCATTTTATGGCTCTGGTCATAACATCAATCTCAATGTGA	868
Qy	292	ProGlnSerSerIleValIlePheLeuLeuPhePheIleTyrGlyLeuSerSerVal	311
Db	869	TTTCATACCTGGCTTCATGGTGATATTACACTCTATAGCTTATAGGCCCTTCTTGATA	928
Qy	312	PhePheAlaLeuMetLeuThrProLeuPheLysLysSerLysHisValGlyIleValGlu	331
Db	929	GCATTGGCTTCCCTCATGAGTGTTTATAAGGAAACCTATGCTCGCTGGTTGGCTGGA	988
Qy	332	PhePheValThrValAlaPheGlyPheIleGlyLeuMetIleIleLeuIleGluSerPhe	351
Db	989	TTTCTCTCACTGATTTTGGGGATGCTGGGA--TTCAGTGGTTATACAGACAACCTT	1045
Qy	352	ProLysSerLeuValTrp-----LeuPheSerProPheCysHisCysThrPheVal	368
Db	1046	CTTTATCTTTGGGATGGATTAAGTCTTCTTAGCCCTTT-----GCCTTCACT	1096
Qy	369	IleGlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSerPheSerAsn	388
Db	1097	GCTGGAATGGCCAGATTACACACTCGATAATTACTTAAGTGGTGTTATTTTCCTGAT	1156
Qy	389	LeuThrAlaGlyProTyrProLeuIleIleThrIleIleMetLeuThrLeuAsnSerIle	408
Db	1157	CCCTCTGGGATTCATACAAATGATAGCCATTTTTCATTTGGCATTTGATACCTTT	1216
Qy	409	PheTyrValLeuLeuAlaValTyrLeuAspGlnValIleProGlyGluPheGlyLeuArg	428
Db	1217	TTCTATTGATATTACATATATTTGAGCGAGTTTACCTGATAAAGATGGCCATGGG	1276
Qy	429	ArgSerSerLeuTyrPheLeuLysProSerTyrTrpSerLysSerLysArgAsnTyrGlu	448
Db	1277	GATTCTCCATTATTTCTTAAATGCTCATTTTGGTCCAAACATCAAAATCACTCATCAT	1336
Qy	449	GluLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGluIleIleGluProVal	468
Db	1337	GAAATCTTGAGATGAATAAATCCCTGAGCATTCCTCTGATGATCTTTTGAACCGGTG	1396
Qy	469	SerSerGluPheValGlyGluAlaIleArgIleSerGlyIleGlnLysThrTyrArg	488
Db	1397	TCTCCAGAAATTCATGGGAAAGAGCCATAGAAATCAGAAATGTTATAAAGAAATATAAT	1456
Qy	489	LysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGln	508
Db	1457	GGAAGACTGGAAGATGAGAGCATTCGACGGCATATTTTTCATATATATGAGGACAG	1516
Qy	509	IleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCys	528
Db	1517	ATCACTGCAATACTTGGGCATAATAGGAGCTGGTAAATCAACACCTGCTAAACATTTCTAGT	1576
Qy	529	GlyLeuCysProProSerAspGlyPheAlaSerIleTyrGlyHisArgValSerGluIle	548
Db	1577	GGATTGCTGTTTCTACAGAGGATTCACCCACTATTATAATACTCAACTCTCTGAATA	1636
Qy	549	AspGluMetPheGluAlaArgLysMetIleGlyIleCysProGlnLeuAspIleHisPhe	568

Db	1637	ACTGACATGGAAGAATAATTAGAAAGATATTGGATTTTGTCCACAGTCAATTTTCATTTT	1691	
Qy	569	AspValLeuThrValGluGluAenLeuSerIleLeuAlaSerIleLeuGlyIleProAla	588	
Db	1697	GACTTCTCACTGTGACAGAGAAAACCTCAGGGTATTTCCTAAAAATAAAGGGATTACGCCA	1756	
Qy	589	AsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLys	608	
Db	1757	AAGCAAGTGGACAAGAGGTAAAGAAATATATAATGGAATTAGACATGCAAGCAATTCAA	1816	
Qy	609	AspAsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAla	628	
Db	1817	GACATATTTCCTAAAAAATTAAGTGTGTGGCAGAGAGAAAACATAACACATAGGGATTGCC	1876	
Qy	629	ValLeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCys	648	
Db	1877	ATCTTAGAGATCTCTCAGGTTTTCCTGCTAGATGACCACTGCTGGATTGGATCCCTTT	1936	
Qy	649	SerArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPhe	668	
Db	1937	TCAAGACACCGAGTGTGGAGCCCTCTGAAGGAGCATAAAGTAGACCGACATATTCCTCTTC	1996	
Qy	669	SerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGln	688	
Db	1997	AGTACCCTCATTCAGATGAGGTGACATCTTGGCTGATAGGAAGATATTCTGCTCTAAT	2056	
Qy	689	GlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyr	708	
Db	2057	GGGAAGTTGAAATGTGCAGATCATCTTTGTTCTGAAGCGAAAGTGGGTATGGATAT	2116	
Qy	709	ArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerLeuValLys	728	
Db	2117	CATTAAAGTTTACACAGCAATGAAATGTGTGACACAGAAAAAATCACATCCCTATTAAAG	2176	
Qy	729	GlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeu	748	
Db	2177	CAGCACATCTCTGATGCCAAGTTAACACAGAAAGTGAAAGAAACCTGTATATAGTTTG	2236	
Qy	749	ProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsn	768	
Db	2237	CCTTTGGAAAAACGAACAAATTTCCAGATCTTTACAGTCACCTTGATAAGTGTCTGAC	2296	
Qy	769	LeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLysLeu	788	
Db	2297	CAGGCGATAAGGAATATGCTTTTCAGTGACATCTCTGAATGAAGTATTTCTTGAACCTA	2356	
Qy	789	GluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGlu	808	
Db	2357	GAAGAAAAATCAGCAATTTGATGAACCAATTTTGACATTTGGAAACAAAGAGAAAAATACAT	2416	
Qy	809	GluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuIleLeuSerGlu	828	
Db	2417	GTGCACAGAATACTCGAGATGATCTCGAATGAACAGGTTCTTTGTTCTCTTCTCTGAA	2476	
Qy	829	ThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGlnMetTyrThrIleAla	848	
Db	2477	ACAGAAGGCT---GTCATGATGTCAGCTCTCTGGACGACGACAAATCTATGCGAGTGCCA	2533	
Qy	849	LysPheHisPhePheThrLeuLysArgGluSerLysSerValArgSerValLeuLeuLeu	868	
Db	2534	ACACTTCGCTTCTTTAAAGTTTAAAGCGGTGAAGGAGAGCTCTTTTGTGTTTGTACTAGTA	2593	
Qy	869	LeuLeuIlePhePheThrValGlnIlePheMetPheLeuValHisHisSerPheLysAsn	888	
Db	2594	CTTGGAAATTTGCTTTT---ATCCCATCATCTACAGAGAATATATGATATAAGTAACCTCGT	2650	
Qy	889	AlaValValProIleLysLeuValProAspLeuTyrPheLeuLysProGlyAspLysPro	908	
Db	2651	GAAACTCATTTGTGGAGTTTTCACCCAGATGATGATTTCCCTTCTCTGGAACAAATCCCG	2710	
Qy	909	HisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeu	928	

Db 2711 AAGAGCCCTCTTACCAGCCTGTTAATCGTTAATAATACAGGATCAAAATATTGAAGACCTC 2770  
 Qy 929 IleserPhePheThrSerGlnAenileMetValThrMet-----IleAen 943  
 Db 2771 GTGCATTCACTAAGTGTGAGATATAGTTTGGAAATAGATGACTTTAGAAACAGAAAT 2830  
 Qy 944 AspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAenValMetHisSerGlu 963  
 Db 2831 GCCTCAGATGATCCCTCC-----TACATGGAGCCATCATAGTGTCTGGTGACCA 2881  
 Qy 964 LysAspTyrValPheAlaAlaValPheAenSerThrMetValTyrSerLeuProIleLeu 983  
 Db 2882 AGGGATTACAGATTCTCTCGGTGTAATACCAAGAAATGAATTGTTTCTGTCTT 2941  
 Qy 984 ValAenileIleSerAsnTyrTyrLeuTyrHisLeuAenValThrGluThrIleGlnIle 1003  
 Db 2942 ATGGGAATTTGTAGCAATGCCCTTATGGGAATTTTAACTTCACGAGCTTATTCAAACG 3001  
 Qy 1004 TrpSerThrProPhePheGlnIleThrAepIleValPheLeyleIleGluLeuTyrPhe 1023  
 Db 3002 GAGAGCACTTCATTTCTCGTGAT-----GACATAGTGTGGATCTTGGTTTATAGAT 3055  
 Qy 1024 GlnAlaAlaLeuGlyIleIleValThrAlaMetProProTyrPheAlaMetGluAen 1043  
 Db 3056 GGTCTCCATATTTTGTGTGATCACAACCTCGCTTCTCTCTTTATCGGCATGAGCAGC 3115  
 Qy 1044 AlaGluAenHisIleLysIleValTyrThrGlnLeuLysLeuSerGlyLeuLeuProSer 1063  
 Db 3116 ATCAGCGATTATAAAAAAATGTTCAATCCAGTTATGGATTTCAGGCCCTCTGGCCTTCA 3175  
 Qy 1064 AlaTyrTrpIleGlyGlnAlaValAlaPheIleProLeuPhePheIleIleLeuLeu 1083  
 Db 3176 GCATACGTGTGGACAGCTCTGGTGACATTCATATATCTTCTGATT----- 3236  
 Qy 1084 MetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPhe--- 1102  
 Db 3227 -----CTCTTTTCAATACATTTAATTTACTACTCATATTTCTGGGATTCAG 3274  
 Qy 1103 -----LeuAlaValValPheCysLeuIleGlyTyrValProSerVal 1116  
 Db 3275 CTTTCATGGAACTCATGTTTGTGTGGTATGATGCAATATGGTGTGCAGTTCTCTT 3334  
 Qy 1117 IleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAenThrLysGluPhe 1136  
 Db 3335 ATATTCCTCACATATGTGCTTCATTCATCTTCGCAAGTGAGAAAAAATAATGGCTTT 3394  
 Qy 1137 TrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhe 1156  
 Db 3395 TGTCCTTTGGCTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3448  
 Qy 1157 PheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIle---IlePro 1175  
 Db 3449 -----ACTCAATATGAAAACTCAACTAAATTTTGTGCATGATTTTCATACCT 3496  
 Qy 1176 IleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAenValarg 1195  
 Db 3497 TCCTTCACCTTTGCTGGGTATGTCATGTTTATTTATTTATTTATTTATTTATTTATTT 3547  
 Qy 1196 LysAenValAspThrTyrAsnProTrpAspArgLeuSer----- 1208  
 Db 3548 AGAACTTGACAGCTGGAC-----AATAGAAATGAAGTCAATAAACCATTCTT 3601  
 Qy 1209 ValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuGlnTyrTyr 1228  
 Db 3602 TTAACAACCTTAATACCATACCTTCAGAGTGTATTTTCTTTTGTCTATAGGTGTCTG 3661  
 Qy 1229 GluLysTyrGlyGlyArgSerIleArgLysAspProPhePheAenLeuSerThr 1248  
 Db 3662 GAATGAAGTATGGAATGAATATGATATGAAGCCCGATTTTCAGA---ATCTCTCCA 3718  
 Qy 1249 LysSerLeuAenArgLysLeuProGluProProAspAenGluAspGluAspVal 1268  
 Db 3719 CGGAGTAGAGAACTCAT---CCCAATCCGGAAGCCCGAAGAAAGATGAAGATGTT 3775

Qy 1269 LysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGluLysPro 1288  
 Db 3776 CAAGCTGAAGAGTCCCAAGCAGCAAAATGCACCTCACTGCTCCAACTTGAGGAGGAACCA 3835  
 Qy 1289 SerIleMetValSerAenLeuHisLysGluTyrAspAspLysLysPheLeuLeuSer 1308  
 Db 3836 GTCATAACTGCAAGCTGTTTACAAAGGAATATTATGAGACAAAGAAAAGTTCTCTTTCA 3895  
 Qy 1309 ArgLysValLysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIle 1328  
 Db 3896 ACAAGAAAGAAAGAAATAGCCATCAGAAATGTTTCTTTTGTGTTAAAAAGTGAAGTT 3955  
 Qy 1329 LeuGlyLeuLeuGlyProAenGlyAlaGlyLysSerThrIleLeuAenIleLeuValGly 1348  
 Db 3956 TTGGGATTACTAGGACACAATGGAGCTGGTAAAGTACTTCCATTAAATGATACTGGG 4015  
 Qy 1349 AspIleGluProThrSerGlyGlnValPheLeu---GlyAspTyrSerSerGluThrSer 1367  
 Db 4016 TGCACAAAGCAACTGAGGAGTGGTGTGTACAAAGGACAGCAGCATCAGTAAGGCAA 4075  
 Qy 1368 GluAspAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAenProLeuTrpPro 1387  
 Db 4076 CAGCATGACAAACAGCCTCAAGTTCTTGGGGTACTGCGCTCAGGAGAACTCCTGTGGCCC 4135  
 Qy 1388 AspThrThrLeuGlnGluHisPheGluIleTyrGlyValAlaValLysGlyMetSerAlaSer 1407  
 Db 4136 AGCTTACATGAAGAGCACTTGGAGTGTATGACAGCTGTGAAGAGCTGGGCAAGAA 4195  
 Qy 1408 AspMetLysGluValIleSerArgIleThrHisAlaLeuAspLysGluHisLeuGln 1427  
 Db 4196 GATGCTGCTCTCAGTATTTTCAAGTTGGTGAAGCTCTTAAAGCTCCAGGAACAACCTTAAG 4255  
 Qy 1428 LysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMet 1447  
 Db 4256 GCTCCTGTGAAAACTCTATCAGAGGGAATAAGAAAGCTGTGCTTTGTGTGACATC 4315  
 Qy 1448 LeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAla 1467  
 Db 4316 CTGGGAACCCATCAGTGGTGTCTTAGATGAGCCGTTACCGGGATGAGCCCGAGGG 4375  
 Qy 1468 LysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAenArgLysArgAlaIle 1487  
 Db 4376 CAGCAGCAATGTGGCAGATCTTACGCTACCGTTAAACAAAGGAGGAGGACCCCTC 4435  
 Qy 1488 LeuThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetVal 1507  
 Db 4436 TTGACCACTTACATGTGCAGAGGCTGAGGCTGTGTGTGACCGTATGCCATGATGGTG 4495  
 Qy 1508 SerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGly 1527  
 Db 4496 TCAGGAACGCTAAGGTGTATTTGTTCCATTCACATCTGAAAAACAAGTTTGTGTAGAT 4555  
 Qy 1528 TyrPheLeuGluIleLysLeuLysAspTrpIleGluAenLeuGluValAspArgLeuGln 1547  
 Db 4556 TATTTACTAGAAATAAATGAAA-----GAACCTACCAGGTGGAGGTCTCCAC 4606  
 Qy 1548 ArgGluIleGlnTyrIlePheProAenAlaSerArgGlnGluSerPheSerIleLeu 1567  
 Db 4607 ACAGAGATTTTGAAGCTTTTCCACAGGCTGTGGCAGGAAAGATATTTCTTTAATG 4666  
 Qy 1568 AlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGlu 1587  
 Db 4667 GCGTATAGTTTACTGTGGAGGATGCCACCTCTATCTCGGCCCTTTTCAAGTTAGAG 4726  
 Qy 1588 GluAlaLysHisAlaPheAlaIleGluTyrSerPheSerGlnIleThrLeuGluGln 1607  
 Db 4727 GCGATGAACAGACCTTCAACCTGGAGGAATACAGCTCTCTCAGGCTACCTTGGAGCAG 4786  
 Qy 1608 ValPheValGluLeuThrLysGluGlnGluAspAsnSerCysGlyThrLeuAen 1627  
 Db 4787 GTATTCTTAGAATCTCTGTAAGAGCAG---GAGCTGGGAAATGTTTATGATGATAAATGAT 4843





Db 847 ACTTTAATGGCTCTTATTGTGTAATAATCTGCACAAATGTGCTCTGACTGGTTTGTGATG 906  
Qy 299 IlePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThr 318  
Db 907 GTCTTCAACCTCTTCTCTCTATGCGCTGCTTTGATAAATTTAGCTTTTCTGATGAGT 966  
Qy 319 ProLeuPheLeuGlySerLysHisValGlyIleValGluPhePheValThrValAlaPhe 338  
Db 967 GTGTGTGATAAGAAACCTTCTTACGGCTGTGTGTGTCTCTCTATTATTGTCTTTGG 1026  
Qy 339 GlyPheIleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTyr-- 357  
Db 1027 GGGATCTGGGATTTCCCA--GCATGTATACACGCTCTCTGCAATTTTGGATGGACT 1083  
Qy 358 -----LeuPheSerProPheCysHisCysThrPheValIleGlyIleAlaGlnValMet 375  
Db 1084 TTTGTGTCTCTTAGGCCCTTT-----GCCTTCACTGTGGGATGGCCCGCTTATA 1134  
Qy 376 HisLeuGlu--AspPheAsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyr 394  
Db 1135 CATTTGGACTATGATGTGAATCTAATGCCACCTGGATCTCT--TCAAAAATCCCATAC 1191  
Qy 395 ProLeuIleIleThrIleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAla 414  
Db 1192 CTCATAATAGTACTCTTTCAATGTGCTTTTGTGACACCTCTCTGTAATTTGGTATTGACA 1251  
Qy 415 ValTyrLeuAspGlnValIleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPhe 434  
Db 1252 TTATATTTTGACAAATTTTGGCCCTCGAATATGGACATCGATGTTCTCCCTCTTTTTC 1311  
Qy 435 LeuLysProSerTyrTyrSerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsn 454  
Db 1312 CTGAAATCTGTGTGTGTTTCAACACGGAAGGCTAATCATGTGTCCTTGAGAATGAA 1371  
Qy 455 ValAsnGlyAsnIleSerPheSerGluIleIleGluProValSerSerGluPheValGly 474  
Db 1372 ACAGATTCTGATCCCTACCACTAATGACTGTTTGAACACGAGTCTCCAGAAATTCGTGGG 1431  
Qy 475 LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnVal 494  
Db 1432 AAAGAAGCAATCAGAAATCAAAATCTTAAAGAAGATATGACGGNAGTGTGAGAGATG 1491  
Qy 495 GluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGly 514  
Db 1492 GAAGCTTTGAAAGGTGTGTGTGTGACATATATGAAGCCAGATCACTGCGCTCTTGGT 1551  
Qy 515 HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysProSer 534  
Db 1552 CACATGGAGCTGGAATAACTCCCTGTAAACATCTTAGTGGTGTGTCAGTTCCAAACA 1611  
Qy 535 AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
Db 1612 TCAGGTTCACTGCTCTATAATCACACACTTTCAAGATGGCTGATAGAAATATC 1671  
Qy 555 ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
Db 1672 AGCAAGTTCACTGGATTTGTGCCAATCCCAATGCAATTTGGATTTCACTGTGAAA 1731  
Qy 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGlu 594  
Db 1732 GAARACCTCAGGCTGTTGCTTAAATAAAGGATTTTGCCACATGAAGTGGAGAAAGAG 1791  
Qy 595 ValGlnLysValLeuLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLys 614  
Db 1792 GTACAACGAGTTGTACAGGAATTAAGAAATGGAATAATTTCAAGCATCTTGTGCTCAAAAC 1851  
Qy 615 LeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLys 634  
Db 1852 TTAAGTGTGGACAAATAGAAACTAATTTGGATTTGCCATTTTAGGAGATCTCAA 1911  
Qy 635 IleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleValTyr 654  
Db 1912 GTTTTGTCTATTGGATGAACCGACTGCTGGATTGGATCTCTTTCAAGGCACCGAATATGG 1971

Qy 655 AsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPheMetAsp 674  
Db 1972 AATCTCTGAAGAGGGGAAATCAGACAGATTAATTTCTCTTCTCAGCACCCAGTTTATAGAT 2031  
Qy 675 GluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLysCysVal 694  
Db 2032 GAGGCTGACATCTTGGCGGACAGGAAGGTGTTTCATATCAATGGGAAGCTGAAGTGTGCA 2091  
Qy 695 GlySerSerMetPheLeuLysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIle 714  
Db 2092 GGTCTCTCTCTGTTCTTAAAGAAATGGGCATAGGTACCATTTAAGTTTGCATCTG 2151  
Qy 715 AspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAla 734  
Db 2152 AATGAAGGTGTGATCCAGAGATTAACATCACTGGTTAAGCAGCACATCTCTGATGCC 2211  
Qy 735 ThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAsp 754  
Db 2212 AATTGACAGCACAAAGTGAAGAAACCTGTATATATTTTGCCTTTGGAAGGACAAAC 2271  
Qy 755 LysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyr 774  
Db 2272 AATTTCCAGAACTTTTACAGGATCTTTGATAGATGTTCTAACCAAGGCATTTGAGGATTAT 2331  
Qy 775 GlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGluIle 794  
Db 2332 GGTGTTTCCATAACAACTTTGAATGAGGTGTTTCTGAAATTAGAAGGAAATCAACTATT 2391  
Qy 795 AspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluGluMetAspSerLys 814  
Db 2392 GATGATCAGATATTGGAATTTGGGACAATTTACAACTGATGGGCAAAAGATATAGGA 2451  
Qy 815 SerPheAspGluMetGlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuVal 834  
Db 2452 AGCTTGTGTGAGCTGGAACAAGTTTGTCTCTTCCACAGAACCAAGGAAACA---ATC 2508  
Qy 835 SerThrMetSerLeuTyrLysGlnGlnMetTyrThrIleAlaLysPheHisPheThr 854  
Db 2509 AGTGGCGTGGCTCTGGAGGAGCAGGCTCTGCAATAGCAAAAGTTTCCTTCTTAAAG 2568  
Qy 855 LeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeuLeuIlePhePheThr 874  
Db 2569 TTAAGAAAGAAAGAAAGAAAGCCCTGTGGACTATATTATTGCTTTTGGTATTAGCTTTATC 2628  
Qy 875 ValGlnIlePheMetPheLeuValHisSerPheLysAsnAlaValValProIleLys 894  
Db 2629 CCTCACTTTTGGACATCTATTCTACGAGTCAATCAGAAAAGT---TACCGTGGGAA 2685  
Qy 895 LeuValProAspLeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSer 914  
Db 2686 CTGCTCTCCAAATACATATCTTCTCTCACGAGCAACCAACCAAGGATCTCTGACCCAT 2745  
Qy 915 LeuLeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
Db 2746 TTACTGTCATCAATAAGACAGGCTCAACCATTTGATTAACATTTTACATCACTGAGGGA 2805  
Qy 935 GlnAsnIleMetValThrMet-----IleAsnAspSerAspTyrValSer 949  
Db 2806 CAGAACATAGCTATAGAGTGGATGCTTGGAACTAGAAATGGCACAGATGACCCATCT 2865  
Qy 950 ValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAla 969  
Db 2866 -----TACAAATGGTCTATCATTTGTGCAGGTGATGAAGATGATCAGATTTTCA 2916  
Qy 970 AlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsn 989  
Db 2917 ATAGCATGTAATACAAAACGGCTGAATTTGCTTCTGCTCTCTGATGTCATTAGCAAT 2976  
Qy 990 TyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTyrSerThrProPhePhe 1009  
Db 2977 GGACTACTTGGAAATTTTAAATTCGTCAGAACACATTCAGACTGACAGAGCACATTTT 3036





Qy	376	HisLeuGlu---AspPheAsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyr	394
Db	1278	CATTGGCATTGATGTGAATCTTAATGGCCACCTGGATTCT---TCACAAAATCCATAC	1334
Qy	395	ProLeuIleIleThrIleIleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAla	414
Db	1335	CTCATATACTACTCTTTTCATGTTGGTTTTTGACACCCCTCTGTATTGGTATTGACA	1394
Qy	415	ValTyrLeuAspGlnValIleProGlyGluPheGlyLeuArgSerLeuTyrPhe	434
Db	1395	TTATATTTTGCACAAAATTTTGGCCGCTGAATATATGCATCATGATGTTCTCCCTTGTTTTTC	1454
Qy	435	LeuLysProSerTyrTrpSerLysSerLysArgAsnTyrGluIleLeuSerGluGlyAsn	454
Db	1455	CTGNAATCTCTGTTTTTGGTTTCAACACGGAGGGCTAATCATCTGGTCTCTTGAGATGAA	1514
Qy	455	ValAsnGlyAsnIleSerPheSerGluIleIleGluProValSerSerGluPheValGly	474
Db	1515	ACAGATTCTGATCTCATCACCTAATGACTGTTTTGAACCACTGTCTCCAGAAATTTCTGTGG	1574
Qy	475	LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnVal	494
Db	1575	AAGAAGCCATCAGATCAAAATCTTAAAAAAGATATGCAGGGAAGTGTGAGAGATG	1634
Qy	495	GluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGly	514
Db	1635	GAAGCTTTGAAAGTGTGGTGTTTGACATATATCAGAGCCAGATCACTGCGCTCCCTGGT	1694
Qy	515	HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysProSer	534
Db	1695	CACAGTGGAGCTGGAAAATCTCCCTGTTTAAACATTTCTTAGTGGTGTGCAGTTCACAA	1754
Qy	535	AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla	554
Db	1755	TCAGGTTTCAGTCACTGTCTTAATCAACACTTTCAAGAATGGCTGATATAGAAAAATATC	1814
Qy	555	ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu	574
Db	1815	AGCAAGTTCAGTGGATTTTGTCCACATCCAAATGTCGAATTTGGATTTCTCATCTGTGAA	1874
Qy	575	GluAsnLeuSerIleLeuAlaSerIleGlyIleProAlaAsnIleIleGlnGlu	594
Db	1875	GAAAACCTCAGCTGTTTGTCTAAATATAAGGGATTTTGCCACATGAAGTGGAGAAAGAG	1934
Qy	595	ValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLys	614
Db	1935	GTACACGAGTGTACAGGAATAGAAATGGAAATATTCAAGACATCTTGTCTCAAAAC	1994
Qy	615	LeuSerGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLys	634
Db	1995	TTAAGTGGTGACAAAATAGAAAATCACTTCTTTGGGATTTGCCATTTTGGAGATCTCTCAA	2054
Qy	635	IleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleValTyr	654
Db	2055	GTTTTGCTATTGGATGAACCACTGCTGGATTGGATCTCTTTCACAGCCACCGAATATGG	2114
Qy	655	AsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPheMetAsp	674
Db	2115	AATCTCTGAAAGAGGGGAAATCAGACAGAGTAATCTCTTTCAGCACCCAGTTTATAGAT	2174
Qy	675	GluAlaAspIleLeuAlaAspArgGlyAlaValIleSerGlnGlyMetLeuLysCysVal	694
Db	2175	GAGCTGACATCTGGCGGACAGAGAGGTGTTTCATATCCAAATGGAGAGCTGAAGTGTGCA	2234
Qy	695	GlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIle	714
Db	2235	GGCTCTTCTGTCTCTTAAAGAAGAAATGGGGCCTAGGCTACCAATTTAAGTTTGCATCTG	2294
Qy	715	AspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAla	734
Db	2295	AATGAAGAGTGTGATCCAGAGAGATATACATCATCTGTTAAGCAGCACATCTCTGATGCC	2354

Qy	735	ThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAsp	754
Db	2355	AAATTCACAGCAAAAGTGAAGAAAACCTGTATATATTTTGCCTTTGGAAAGCACAAC	2414
Qy	755	LysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyr	774
Db	2415	AAATTTCCAGAACTTTTACAGGGATCTGTAGATGTTCTAAACCAAGCATTTAGGATTAT	2474
Qy	775	GlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGlnValGluAlaGluIle	794
Db	2475	GGTGTTCCTACAACTTTGAATGAGGTGTTTCTGAAATTAGAGGAAATCAACTATT	2534
Qy	795	AspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluMetAspSerLys	814
Db	2535	GATGAATCAGATATTGGAATTTGGGACAAATTACAACATGATGGGCAAAAGATATAGGA	2594
Qy	815	SerPheAspGluMetGluGlnSerLeuLeuLeuSerGluThrLysAlaSerLeuVal	834
Db	2595	AGCTTGTGTGAGTGGACAAAGTTTGTCTCTCTCCACGAACAACAGGAAAAACA--ATC	2651
Qy	835	SerThrMetSerLeuTyrLysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThr	854
Db	2652	AGTGGGTGGCGCTCTGGAGGCACAGGCTGTGTGCAATAGCAAAAGTTGCCTTCTTAAG	2711
Qy	855	LeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeuLeuIlePhePheThr	874
Db	2712	TTAAAGAAAGAAAGAAAGCCGTGGACATATATTTGCTTTTGGTATTAGCTTTATC	2771
Qy	875	ValGlnIlePheMetPheLeuValHisHisSerPheLysAsnAlaValValProIleLys	894
Db	2772	CCTCAACTTTTGAACAATCTATTCTACGAGTCATATCAGAAAAAGT---TACCCTGGGAA	2828
Qy	895	LeuValProAspLeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSer	914
Db	2829	CTGTCTCCAAATACATATCTTCTCTCACAGSACACACACAGGATCCTCTGACCCAT	2888
Qy	915	LeuLeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeuIleSerPhePheThrSer	934
Db	2889	TTACTGTCTCAATATAGACAGGTCACCACTTGATGAACCTTTTTCACATCAGTGAGCGGA	2948
Qy	935	GlnAsnIleMetValThrMet-----IleAsnAspSerAspTyrValSer	949
Db	2949	CAGAACATAGCTATAGAAGTGGATGCCTTTGGAACTAGAAATGGCAGATGACCCATCT	3008
Qy	950	ValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAla	969
Db	3009	-----TACAAATGTGTCTATCTTGTCTCAGGTGATGAAAGAGGATCACAGATTTC	3059
Qy	970	AlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIleLeuSerAsn	989
Db	3060	ATAGCATGTAAACAAAACGGCTGAATGCTTTCTCTCTCTGGATGTCAATAGCAAT	3119
Qy	990	TyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleThrPheThrProPhePhe	1009
Db	3120	GGACTACTTGGAAATTTTAAATTCGTACAGACACATTCAGACTCAGACAGACACATTTT	3179
Qy	1010	GlnGluIleThrAspIleVal-----PheLysIleGluLeuTyrPheGlnAlaLeu	1027
Db	3180	GAAGGCATATGGATTATGAGTATGGGTACCGAAGTAACACCTTCTCTGGATACCGAT	3239
Qy	1028	LeuGlyIleIleValThrAlaMetProProTyrPheAlaMetGluAsnAlaGluAsnHis	1047
Db	3240	-----GCAGCGCTCTTTCACCTCCATACATTCGAATGACAGCAGATTTGGTACTAC	3287
Qy	1048	LysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIle	1067
Db	3288	AAAAAAAAGCTATCTCCAGCTACGGATTTACGGCCTCTACCCCTCTCTGCATACTGGT	3347
Qy	1068	GlyGlnAlaValValAspIleProLeuPhePheIleIleLeuIleLeuMetLeuGlySer	1088
Db	3348	GGCCAAGCAGCTGGTGGATGTTTCCCTGTACTTTTGTATCTCTCTCTGCTAAATG----	3407
Qy	1088	LeuLeuAlaPheHisTyrGly-----LeuTyrPheTyrThrValLysPheLeuAla	1104

Db	3402	ATAATGATATATATTTTATGCCCAGAGAGATATATTTATTAATTTAAAAACCTGTTAATT	3461
Qy	1105	ValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSer	1124
Db	3462	CAAAATCCTGTGTAGTATTGGCTATGTCTCATCTCTCTTTTCTTGACATATGTAATTC	3521
Qy	1125	PheThrPheIlyslsIleLeuAsnThrLysGluPheThrPheSerPheIleTyrSerValAla	1144
Db	3522	TTTCATTTTTCCGAATGGAGAAAATAAGTGGCAATTTGGTCATTTTTTCTTTCTTAATGTG	3581
Qy	1145	AlaLeuAlaCysIleAlaIleThrGluIleThr-----PhePheMetGlyTyrThrIle	1162
Db	3582	GTCACTTCTCGATAGTTGCTACTGACTAAATGAATATGGAATTTCTAGGGCTATTTTTT	3641
Qy	1163	AlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCys	1182
Db	3642	GGCACCATGTTA-----ATACCTCCCTTCATTTGATGGTCTCT	3680
Qy	1183	LeuIleSerPheIleIlyslsSerTrpLysAsnValArgLysAsnValAspThrTyrAsn	1202
Db	3681	CTATTCTATTTTTCTGAGATT-----	3701
Qy	1203	ProTrpAspArgLeuSer-----ValAlaVal	1211
Db	3702	CCTCCTGATTCCTGGATTCCTTAGGAGCTTCAGAATCTCAAAATGTATACCTGGCACGT	3761
Qy	1212	IleSerProTyrIleuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysLys	1231
Db	3762	CTGATCCTTACCTTCATTTTCTCATTTTCTTTTTCATTCGCGATCGCTAGAAATGAAC	3821
Qy	1232	TyrGlyGlyArgSerIleArgLysAspProPheArgAsnLeuSerThrLysSerLys	1251
Db	3822	TGCAGGAAGAACTAATGAGAAGGATCCTGTGTTCAGA---ATTTCCTCAAGAAGC---	3875
Qy	1252	AsnArgLysLeuProGluProProAspAsnGluAspGluAspValLysAlaGlu	1271
Db	3876	AAGCCTATTTTCCAAACCCAGAGAGCCTGGAAGGAGAGGAGGAAGATATCCAGATGGAA	3935
Qy	1272	ArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGlnLysProSerIleMet	1291
Db	3936	AGAATGAGACAGTGAATGCTATGGCTGTCGAGACTTGTATGAGACACCCGCTCATATT	3995
Qy	1292	ValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSerArgLysVal	1311
Db	3996	GCCAGCTGTCTACGGAAGGAATATGCAGGCCAAAAGAAAAATTTGCTTTCTTAAAGGAAG	4055
Qy	1312	LysLysValAlaThrIlyslsTyrIleSerPheCysValLysLysGlyGluIleLeuGlyLeu	1331
Db	4056	AAACAAATTTCCACAGAAATGTCTCTTTTGTGTAAAGGTTAAAGGTGAAGTTATAGACTG	4115
Qy	1332	LeuGlyProAsnGlyAlaGlyLysSerThrIleIleAsnIleLeuValGlyAspIleGlu	1351
Db	4116	TTAGGACACAATGGGCTGTGTAAGTACAACTATTAAAGTATGATACTGGAGACACAAA	4175
Qy	1352	ProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAspAsp	1371
Db	4176	CCAACCTGCAGGACAGGTGATTTTG-----AAAGGGAGCGGTGAGGGGAA	4220
Qy	1372	SerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTrpProAspThrThrLeu	1391
Db	4221	CCCTGGGCTTCTGGGTACTGCCCTCAGGAGAAATGGCGTGTGGCCCACTTCAGAGTG	4280
Qy	1392	GlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGlu	1411
Db	4281	AGGCAGCACCTGGAGGTGATCGCTGCTCGGTGAAAGGTCTCAGGAAGGGGACGCAATGATC	4340
Qy	1412	ValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisIleuGlnLysThrValLys	1431
Db	4341	GCCATCACAGGTTAGTGATGTCGCTCAAGCTCGAGGACGAGCTGGAAGGCTCCCGTGAAG	4400
Qy	1432	LysLeuProAlaGlyIleLysArgLysLeuLysCysPheAlaLeuSerMetLeuGlyAsnPro	1451

Db	4401	ACCTTGTGAGAGGAATAAAGCAAAAGCTGTGCTTTGTCTGAGCANTCTCGGGAACCCG	4460
Qy	1452	GlnIleThrLeuLeuAspGluProSerThrGlyMetAspProClylsAlaLysGlnHisMet	1471
Db	4461	TCAGTGTGTCTTCTGATGAGCCGTGCACGGGATGACCCCGAGGGCAGCAGCAATG	4520
Qy	1472	TrpArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeuThrHis	1491
Db	4521	TGCAGGTGATTGGGCCACCTTTAGAAACACGAGAGGGGGCCCTCTCTGACCACCCAC	4580
Qy	1492	TyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeu	1511
Db	4581	TACATGCAGAGCTCAGGCGGTGTGTGACCGAGTGGCCATCATGTGTTCAGGAAGCTG	4640
Qy	1512	ArgCysIleGlyThrValGlnHisIleuLysSerLysPheGlyLysGlyTyrPheLeuGlu	1531
Db	4641	AGATGATTGTGTCCATCAACACCTCGAAAGCAAATTTGGCAAGAGACTCTCTGGGAG	4700
Qy	1532	IleLysLeuLysAspTrpIleGluAenLeu---GluValAspArgLeuGlnArgGluIle	1550
Db	4701	ATGAAGCTGAAG-----AACCTGGCACAAATGAGGCCCTCCATGCAGAGATC	4748
Qy	1551	GlnTyrIlePheProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAlaTyrLys	1570
Db	4749	CTGAGGCTTTTCCCCCAGAGTGCTCAGCAGGAAGGTTCTCTCTCGTAGGTCTATAAG	4808
Qy	1571	IleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluAlaLys	1590
Db	4809	TTCCCTGTTCAGGATGTGGACCTTTATCACAAGGCTTCTTCAAATTTAGAGATAGTAAA	4868
Qy	1591	HisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheVal	1610
Db	4869	CAGAGTTTCACCTTCAGGAGGATACACCTCTCACAGTCTTACCCTTGAGCAGGTTTTCCCG	4928
Qy	1611	GluLeuThrLysGluGlnGlu-----GluGluAsp	1620
Db	4929	GAGCTCTCAAGAGAGCAGAGAGTGGGTGATCTTTGAAGAGGAC	4970

Search completed: April 12, 2004, 21:51:58  
Job time : 1315 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 12, 2004, 21:08:44 ; Search time 176 Seconds  
(without alignments)  
5177.440 Million cell updates/sec

Title: US-10-090-458-5

Perfect score: 8426

Sequence: 1 MSTAIREVGVRQRTILLK.....CGTLNLSILWERTQEDRVVF 1642

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/usfto\_spool/US10090458/runat\_08042004\_121553\_25554/app\_query.fasta\_1.1799  
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10090458 @cgn1.1.6 @runat\_08042004\_121553\_25554 -NCPUs=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT\_DSBLCK=100 -LONGLOG  
-FMT\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1:	/cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2:	/cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3:	/cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4:	/cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5:	/cgn2_6/ptodata/2/ina/FACTUS.COMB.seq:*
6:	/cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553.5	18.4	6525	US-08-762-500-74	Sequence 74, Appl
2	1521.5	18.1	5894	US-08-665-259-24	Sequence 24, Appl
3	1521.5	18.1	5894	US-08-762-500-24	Sequence 24, Appl
4	1479	17.6	7860	US-09-526-193A-2	Sequence 2, Appl
5	1248	14.8	974	US-09-833-381-1085	Sequence 1085, Ap
6	900.5	10.7	1350	US-09-833-381-1082	Sequence 1082, Ap
7	647	7.7	3609	US-09-833-381-958	Sequence 958, App
8	535	6.1	485	US-09-621-976-409	Sequence 409, App
9	467	5.5	2075	US-09-833-381-962	Sequence 962, App
10	461	5.5	1770	US-09-328-352-3466	Sequence 3466, Ap
11	426	5.1	1767	US-09-489-039A-6408	Sequence 6408, Ap
12	421.5	5.0	1824	US-09-253-991A-1780	Sequence 1780, Ap

13	413.5	4.9	15567	4	US-09-627-376-3	Sequence 3, Appl
14	407.5	4.8	4233	3	US-09-120-513-1	Sequence 1, Appl
15	407.5	4.8	4233	3	US-09-450-105-1	Sequence 1, Appl
16	404.5	4.8	4186	4	US-09-672-810-1	Sequence 1, Appl
17	404	4.8	1785	4	US-09-543-681A-1356	Sequence 1356, Ap
18	402	4.8	4195	4	US-09-672-810-3	Sequence 3, Appl
19	400.5	4.8	4646	1	US-08-181-471-2	Sequence 2, Appl
20	400.5	4.8	4646	1	US-09-023-655-1167	Sequence 1167, Ap
21	398.5	4.8	4669	6	5206352-3	Patent No. 5206352
22	398.5	4.7	4264	2	US-08-784-649A-1	Sequence 1, Appl
23	398.5	4.7	4264	2	US-08-784-649A-5	Sequence 5, Appl
24	396.5	4.7	6505	2	US-08-793-610-5	Sequence 5, Appl
25	396.5	4.7	8630	4	US-09-306-417-1	Sequence 1, Appl
26	396.5	4.7	8630	4	US-09-306-417-2	Sequence 2, Appl
27	396.5	4.7	9318	2	US-08-793-610-6	Sequence 6, Appl
28	395.5	4.7	4669	2	US-08-752-447-1	Sequence 1, Appl
29	395.5	4.7	4669	4	US-09-316-167-1	Sequence 1, Appl
30	395.5	4.7	4669	4	US-09-397-233-1	Sequence 1, Appl
31	395	4.7	3988	4	US-09-762-195-1	Sequence 1, Appl
32	394	4.7	4669	2	US-08-583-276-18	Sequence 18, Appl
33	392	4.7	14672	4	US-08-961-527-111	Sequence 111, App
34	379.5	4.5	580073	4	US-08-545-528D-1	Sequence 1, Appl
35	368	4.4	6130	2	US-08-604-488-1	Sequence 1, Appl
36	368	4.4	6130	3	US-08-030-081-1	Sequence 1168, Ap
37	366	4.3	3924	4	US-09-023-655-1168	Sequence 2, Appl
38	366	4.3	3924	4	US-09-762-195-2	Sequence 3, Appl
39	365	4.3	4443	4	US-09-425-453A-3	Sequence 9, Appl
40	365	4.3	4443	4	US-09-425-453A-9	Sequence 17, Appl
41	365	4.3	4443	4	US-09-425-453A-17	Sequence 1, Appl
42	365	4.3	4560	3	US-09-256-703-1	Sequence 1, Appl
43	365	4.3	6129	1	US-07-637-621-1	Sequence 1, Appl
44	365	4.3	6129	2	US-08-951-912-1	Sequence 1, Appl
45	365	4.3	6129	3	US-08-681-838A-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-08-762-500-74  
; Sequence 74, Application US/087625500  
; Patent No. 6030806  
; GENERAL INFORMATION:  
; APPLICANT: Landes, Gregory M.  
; APPLICANT: Burn, Timothy C.  
; APPLICANT: Connors, Timothy D.  
; APPLICANT: Dackowski, William R.  
; APPLICANT: Van Raay, Terence J.  
; APPLICANT: Klinger, Katherine W.  
; TITLE OF INVENTION: NOVEL HUMAN CHROMOSOME 16 GENES, AND USING SAME  
; NUMBER OF SEQUENCES: 83  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENZYME CORPORATION  
; STREET: One Mountain Road  
; CITY: Framingham  
; STATE: Massachusetts  
; COUNTRY: United States of America  
; ZIP: 01701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/762,500  
; FILING DATE: 09-DEC-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/665,259  
; FILING DATE: 17-JUN-1996  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/10469  
; FILING DATE: 17-JUN-1996





Db	2187	TTTCAGGTTGGAAATAAGGACAGCGCGCGCTGACAGACCTGCAACCTCAACCTGTACGAG	2246
Qy	507	GlyGlnIleThrAlaLeuLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIle	526
Db	2247	GGACAGATCACCGTCTCTGGGCCAACACGGTCCGGGAAGACCAACCCCTCTCCATG	2306
Qy	527	LeuCysGlyLeuCysProProSerAspGlyPheAlaSerIleTyrGlyHisArgValSer	546
Db	2307	CTCACAGGTCTCTTTCCCGCCACACAGTGGACGGGCATACATACAGCGGGTATGAATTTCC	2366
Qy	547	GlulAlaAspGluMetPheGluAlaArgLysMetIleGlyIleCysProGlnLeuAspIle	566
Db	2367	CAG-----GACATGGTTCAGATCCGGAAGACCTGGCGCTGTGCCCGACGACGACATC	2420
Qy	567	HisPheAspValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIleLysGlyIle	586
Db	2421	CTGTGTTTCACAACTTGACAGTGCAGACGACACCTTTATTTCACGCCACAGCTGGAAGGGCGTG	2480
Qy	587	ProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAspMetGlnThr	606
Db	2481	TCAGCTCAGAGTGCCTCGAAGAGTCAAGCAGATGCTGCACATCATCGCGCTGGAGGAC	2540
Qy	607	IleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGly	626
Db	2541	AAGTGGAACTCACGAGCGCTTCTCAGCGGGGCGCATGAGGCGCAAGCTCTCCATCGCG	2600
Qy	627	IleAlaValLeuGlyAsnProLysLleLeuLeuLeuAspGluProThrAlaGlyMetAsp	646
Db	2601	ATCGCCCTCATCGGAGGCTCCAGGTGCTGATCTGGACGAGGCCACCTCGGGCATGGAC	2660
Qy	647	ProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsnArgValThr	666
Db	2661	GCCATCTCAGGAGGGCCATCTGGGATCTCTTCAGCGCGCAGAAAAAGTACCGCACCATC	2720
Qy	667	ValPheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIle	686
Db	2721	GTGTGACCAACCCACTTCATGGACGAGGCTGACCTGCTGGGAGCCGCATCGCCATCATG	2780
Qy	687	SerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLysTrpGlyIle	706
Db	2781	GCCAAGGGGAGCTGCAGTGTGCGGGTCTCTCGTGTCTCTCAAGCAGAATACGGTGCC	2840
Qy	707	GlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerSerLeu	726
Db	2841	GGCTATCACATGACGTGGTGAAGGACCGCAGCTGCAACCGGGAAGACATCTCCAGCTG	2900
Qy	727	ValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnLeuValTyr	746
Db	2901	GTCCACCACACGTGCCCAACGCCACCGCTGGAGAGACGCGTGGGGCGGAGTGTCTTTT	2960
Qy	747	SerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeuAsp---Ser	765
Db	2961	ATCCTTCCAGAGAGACGACGACAGGTTTGAAGGTCTCTTTGTAAACTGGAGAAGAAG	3020
Qy	766	HisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPhe	785
Db	3021	CAGAAAGAGCTGGGCATGTCAGCTTTGGGGCATCCATCCACCATGGAGGAAGTCTTC	3080
Qy	786	LeuLys-----	797
Db	3081	CTTCGGTGGGAAGCTGGTGACAGCAGATGACATCCAGGCCATCCAGCTCCCTGCC	3140
Qy	788	LeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeu	807
Db	3141	CTGCAGTACCAACGACGAGAGCGCGCAGGGACTGGGGCTGTG-----	3182
Qy	808	GluGluGluMetAspSerLysSerPheAspGluMetGluGlnSer-----	822
Db	3183	-----GACAGCAACCTCTGTGGGGCCATGACCCCTCCGACGGCATGGAGGCC	3230
Qy	823	LeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGln	842

3231	CTCATGAGAGAGAGCCACCGCTGTCTCAAGCTCAA	CACTGGGGCTCGCCCTGCACCTGCCAG	3239
843	GlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLys		857
3291	CAATTCTGGGCCATG	TTCTGGAAGAGCGGCATACAGCTGGCGC	3335
858	GluSerLysSerValArgSer	ValLeuLeuLeuLeu	868
3336	GAGTGGAAATAGTGGCGGCACAGGTCCTGGTGCCTCTGA	CTGACCTGGCCCTC	3395
869	LeuLeuIlePhePheThrValGlnIlePhe	MetPheLeuValHis	884
3396	CTGCCATCAACTCTCTCGGAGCTCTTCGACGACCCCATGCT	GAGGCTGACCTTGGC	3455
885	SerPheLysAsnAlaValValProIleLysLeuValPro	AspLeuTyrPheLeuLysPro	904
3456	GAGTACGGCAGAACCCGCTGGGCC	TTCTCAGTTCC	3491
905	GlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer	AlaAspSerAsp	922
3492	GGGACCTCCAGCTGGGTCTGACGACCTGTCTGAAGAGC	CACTGCAGGCTGAG	3551
925	IleSerAspLeuSerPhe		931
3552	GGACAGAGAGCCCGCGAGGTGCTCGGTGACCTCGAGGAGT	CTTGATCTTCAGGGCTTCT	3611
932	PheThrSerGlnAsnIleMetValThrMetIleAsnAspSer	Asp	946
3612	GTGAGGGGGCGGCTTTANTAGAGGGTGCTTGTGGCAGCG	TCCTTCAGAT	3665
947	TyrValSerValAlaProHisSerAlaAlaLeuAsnVal	MethHisSerGluLysAspTyr	966
3666	GTGGAGAGCGCACGGTCGTCAACGGCTTGTTCACAC	CAACCAAGGGGTAC	3713
967	ValPheAlaAlaValPheAsnSerThrMetValTyrSer	Leu	980
3714	CACCTCCAGCCACTGCCCTGGCGCTGTGGACAC	CTTCGTTCAGCTGCTG	3767
981	ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHis	LeuAsnValThr	998
3768	TGGGGCTCACGCCTCACTTGTGGTCTCAACTTC	CCACAGCCCGG	3815
999	GluThrIleGlnIleTyrSerThrProPhePheGlnGluIle	ThrAspIleValPheLys	1018
3816	AGCGCCCTGACGCTGCCAAGACCAGTAAACGAGGCGCG	GAAGGATTCGACATGGCC	3875
1019	IleGluLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleVal	ThrAlaMetProTyr	1038
3876	CTAACCTGTCTCTGGCCATGGCATCTTGGCC	AGCACG	3914
1039	PheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyr	ThrGlnLeuLysLeuSer	1058
3915	TTCTCCATCTGCGCGTTCACGAGAGCGCGCTGCAGGCC	CAAGCATGTGCAGTTGTGAGT	3974
1059	GlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaVal	ValAspIleProLeuPhePhe	1078
3975	GGAGTCACGTGGCAGATTCTGGCTCTGCTCTGCTGTG	GGACCTCATCTCTCTC	4034
1079	IleIleLeuIleLeuMetLeuGlySerLeuLeuAlaPheHis	TyrGlyLeuTyrPheTyr	1098
4035	ATCCCACTGTCTGTCTGTGTGTGTAAAGCCCTTCAC	GTGGCTTCAGCGCG	4094
1099	ThrValLysPheLeuAlaValValPheCysLeuIleGlyTyr	ValProSerValIleLeu	1118
4095	GACGGCCACATGGCTGACCCCTGCTGCTCTCTGTCT	TACGGTGGGCCATCATCC	4154
1119	PheThrTyrIleAlaSerPheThrPhe	LysLysIleLeuAsnThrLysGlu	1135
4155	CTCATGTACCTGATGAATCTTCTTCTTGGGGCGGCC	CACTGCCATACAGAGGCTGACC	4214
1136	PheTrpSerPheIleTyrSerValAlaAlaLeuAlaCys	IleAlaIleThrGluIleThr	1155
4215	ATCTTCAACATCTGTGACGATCGCCACTTCTGTGATGT	GTACCATCATGCGCATCCCA	4274



```

Qy 1185 SerPheIleIysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnProTrp 1204
Db 844 -----TTACATTTTGAAATCGTGAATTTTCGAAATAATTTTTCGAGCCAGAA 891
Qy 1205 AspArgLeuSer---ValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePhe 1223
Db 892 AGGGAGTAATCCATCTGTTTTCCTAGTCCCTCTTTCATTTTGGCCATCTTCGTAGTT 951
Qy 1224 LeuLeuGlnTyrTyrGluLysIleTyrGlyArgSerIleArgLysAspProPhePhe 1243
Db 952 ATTCCTCGATGATGAAGGAAGTTTGAATGAACCAATAGAACGGATCTCTGCTTT 1011
Qy 1244 ArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAsp 1263
Db 1012 AGA---ATTTCCTCAAGAGC---GATCGTGTATTTAACAATCCAGAGACCCAGATGGG 1065
Qy 1264 GluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetClyCysGlnCys 1283
Db 1066 GAAGATGAAGATGTAGTCAGGAAGAGTGGAGAGGAGTGGACAGCAAAATGCTTTTGACCTCCGCCGAC 1125
Qy 1284 CysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspLysLys 1303
Db 1126 TTCAGGAGAGCCAGCCATCATTTGCCAGTCCCTTCGCAAGGAGGTACAAAGGAAAGAAG 1185
Qy 1304 AspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCysVal 1323
Db 1186 AATATGCTTTTCTTAAGCAAGAAAGATTCGCAAGGAGTGGTAAAGATCATCCATT 1245
Qy 1324 LysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValaGlyLysSerThrIleIle 1343
Db 1246 CGAAAGGTGAGGTGTAGATTAAGTGGACACATGGAGTGGTAAAGATCATCCATT 1305
Qy 1344 AsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSer 1363
Db 1306 AAGATGATACTGGAGAAACCAAGCCAGCGCGGACAGTGTCTGCTG----- 1353
Qy 1364 SerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsn 1383
Db 1354 ---AAGGGAGACAGCAGGGGACACCCCGGGTTCCTGGGTACTGTCTCAGGAAT 1410
Qy 1384 ProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGly 1403
Db 1411 GCTGTGTGGTCAACCTGACAGTGGAGGAAACACCTGGAGATATTTGCAGCCATAAAGGGG 1470
Qy 1404 MetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLys 1423
Db 1471 ATGAAAAAAGTGACGGCAACGTCGGCATCGAAAGGTGGCAGATCGGCTCAAGCTGCAG 1530
Qy 1424 GluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPhe 1443
Db 1531 GACCAGCTGAAGTCTCCTGTGAAGACGCTGTACAGAGGAGTGAAGAGAAAGCTGTCTTT 1590
Qy 1444 AlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMet 1463
Db 1591 GTGCTGAGCATCTCGGGGAACCTTCGTGTGTACTTCTGTATGAGCCATCCACCGGATG 1650
Qy 1464 AspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLys 1483
Db 1651 GACCCGAGGGGACAGCAAAATGTGGAGCTATCCAGGCCACCTTTTCAAAACACAGAA 1710
Qy 1484 ArgAlaAlaIleLeuThrHisTyrMetGluGluAlaValCysAspArgVal 1503
Db 1711 CGGGCGCCCTCTCGACACCATTCATGCGAGAGCTGAGGCTGTGTGTGACCGGTG 1770
Qy 1504 AlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLys 1523
Db 1771 GCCATCATGTGTCTGGAGGCTGAGATGATTTGGATCCATCCAAATCTGAAGACAAA 1830
Qy 1524 PheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeuGluVal 1543
Db 1831 TTCGCAAAAGAAATATCTCTGGAGATGAAGTGAAG-----ACCCCTCGAGGTG 1881
Qy 1544 AspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSerPhe 1563

```

Search completed: April 13, 2004, 02:03:54  
Job time : 6005 secs

```

Db 1882 GAGCCCTTGAACACGAGGATCATGAGGCTGTTCCCCAGGCTGCTCGGAGAAAGGTAC 1941
Qy 1564 SerSerIleLeuAlaTyrIysIleProLysGluAspValGlnSerLeuSerGlnSerPhe 1583
Db 1942 TCTTCTGTATGTCTATAAATGCCCAGGGAAGATGTGCAACCTTTATCCAGGCTTC 2001
Qy 1584 PheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAla 1603
Db 2002 TTCAATATTAGAGACAGTTAAACAGAGCTTTGACCTGGAGGAGTACAGCCTTTTCACAGTCC 2061
Qy 1604 ThrLeuGluGlnValPheValGluLeuThrLysGluGluGlu-----GluGlu 1619
Db 2062 ACCCTGGAGCAGGCTCTCTCTGGAGCTCTTAAGGAGCAGGAGTGGACCGGTTTGAAGAG 2121
Qy 1620 AspAsnSerCysGlyThrLeuAsnSerThrLeuTrpGlu 1633
Db 2122 GAA-----CTTGATCCCTCAGTCAAGTGGAAAG 2148

```



QY 1218 CysValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIle 1237  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1238 ArgLysAspProPhePhe-----ArgAsnLeuSerThrLysSerLysAsnArgLys 1254  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1251 ---CCACATTTTGGCTGGGACGAGGCTCATCGACATGGTGAAAAACAGGCA 5301  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1255 LeuProGlu----- 1257  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5302 ATGCGCTGATGNNNTGGAAGGTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5361  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1257 ----- 1257  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5362 NNN 5421  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1258 -----ProPro 1259  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5422 NNN 5481  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1260 AspAsnGluAspGluAspGluValLysAlaGluArgLeuLysValLysGluLeuMet 1279  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5482 CTGAAT---GATGAAGATGAAGATGTGAGCGGGAAGACAGAGATTTCTTGATGGTGA 5538  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1280 GlyCysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyr 1299  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5539 GGC-----CAGAATGATATCTTAGAATCAAGGAGTTGACGAAGATATAT 5583  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1300 AspAspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIle 1319  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5584 -----AGAAGGAGCGGAAGCTGCTGTGTGACAGGAT 5616  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1320 SerPheCysValLysLysGlyGluLeuLeuGlyProAsnGlyAlaGlyLys 1339  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5617 TCGGTGGCATCTCTCTGGTGGTGGTCTCTGGGAGTTATGGGGCTGGAAA 5676  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1340 SerThrIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeu 1359  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5677 TCATCNCNTTCAAGATGTTAACAGGAGATACCACTGTTACCNAGGAGATGCTTCTCT 5736  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1360 GlyAspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCys 1379  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5737 ---AACANNNNNTATCTTATCAACATCCATGAGTACATCAGACATGGGCTACTGC 5793  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1380 ProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluLysGly 1399  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5794 CTTGAGGAGGAGTCCAG 5853  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1400 AlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAla 1419  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5854 CTTTGGAGGAGGAGTCCAG 5913  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1420 LeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyLysArg 1439  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5914 CTGGGCTCGTGAAGTATGGAG 5973  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1440 LysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluPro 1459  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 5974 AAGCTCTCTACAGCCATGCTTTGATCGCGGGCTCTCTGGTGGTGGTGGTGGTGGTGGTGG 6033  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1460 SerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPhe 1479  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6034 ACCACAGGATGGATCCCAAGCCCGGGTCTTGTGGAATTTGTGCCCTAAGTGTGTGC 6093  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1480 LysAsnArgLysArgAlaIleLeuThrHisTyrMetGluGluAlaGluAlaVal 1499  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6094 AAG---GAGGGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6150  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1500 CysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHis 1519  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6151 TGCCTAGGATGGCAATCATGTCATGAGAGAGTTCAGTCCCTGGCAGTGTCCAGCAT 6210

QY 1520 LeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGlu 1539  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6211 CTAATAAATAGGTTGGAGATGGTTATACATATTTGTACGAATAGCAGG-----TCC 6264  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1540 AsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArg 1559  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6265 AACCGGACCTGAAGCTGTCCAGGATTTCTTGGACTTGCATTTCTCGAAGTGTCTA 6324  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1560 GlnGluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeu 1579  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6325 AAAGAGAAACACCGGACATCTCAATACAGCTTCCA-----TCTCATTTCTTCTCTG 6381  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1580 SerGlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSer 1599  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6382 GCCAGATATTCAGCATCTCTCCAGAGCAAAAGGAGCTCCACATAGAAGACTACTCT 6441  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1600 PheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGlu 1619  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6442 GTTCTCAGCAACACTTGCACCAAGTATTTGTGAATTTGCCAAGGACCAAGATGATGAT 6501  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1620 AspAsn-----SerCysGlyThrLeu 1626  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6502 GACCACTTAAGGACCTCTCATACAAAAACACAGACAGTAGTGGAGTTGCAGTTCTC 6561  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 1627 AsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640  
 DB TCTGTGTTCTTGAATCTTC----- 5250  
 QY 6562 ACATCTTTTCTACAGGATGAGAAAGTGAAGAAAGCTATGTA 6603

RESULT 15  
 AK031843 2844 bp mRNA linear HTC 18-SEP-2003  
 LOCUS Mus musculus adult male medulla oblongata cDNA, RIKEN full-length  
 DEFINITION enriched library, clone:630417All product:ATP-BINDING CASSETTE  
 PROTEIN homolog [Mus musculus], full insert sequence.  
 ACCESSION AK031843.1 GI:26327664  
 VERSION AK031843  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159  
 3  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,  
 Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multiplexed sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 4  
 AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the  
 FANTOM Consortium.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 JOURNAL Nature 409, 685-690 (2001)







```

Db 6286 CAGGATTTCTTTGGACTTGCATTTCCTGGAAGTGTCTTAAAGAGAAACACCGGAACATG 6345
Qy 1567 LeuAlaTyLysileProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeu 1586
Db 6346 CTACATACAGCTTCA---TCCTCATATCTCTCTGGCCAGGATATTCAGCATCCTC 6402
Qy 1587 GluGluAlaLysHisAlaPheAlaileGluGluTyrsPheSerGlnAlaThrLeuGlu 1606
Db 6403 TCCAGAGCAAAAGCGACTCCACATAGAGACTCTCTGTTCTCAGACAACACACTTGAC 6462
Qy 1607 GlnValPheValGluLeuThrLysGluGlnGluGluGluAspAsn----- 1621
Db 6463 CAAGTATTTGTGAATTTGCCAAGGACCAAGTGAATGATGACCACTTAAAGACCTTCA 6522
Qy 1622 -----SerCysGlyThrLeuAsnSerThrLeuTrpTrpGlu 1633
Db 6523 TTACACAAAACACAGACAGTAGTGAGCTTGCGAGTCTTCACATCTTTTCTACAGGATGAG 6582
Qy 1634 ArgThrGlnGluAspArgVal 1640
Db 6583 AAAGTGAAGAAAGCTATGTA 6603

```

RESULT 13  
CD250953 850 bp mRNA linear EST 22-MAY-2003  
LOCUS AGENCOURT 14214041 NIH\_MGC\_179 Homo sapiens cDNA clone  
DEFINITION IMAGE:30385589 5', mRNA sequence.

ACCESSION CD250953  
VERSION CD250953.1 GI:31011419  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1. (bases 1 to 850)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: [gcaps-remail.nih.gov](mailto:gcaps-remail.nih.gov)  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: NDAM452 row: 0 column: 06  
High quality sequence stop: 719.  
Location/Qualifiers

FEATURES  
source  
1..850  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30385589"  
/tissue\_type="Pituitary"  
/lab\_hosts="DHI08-Ton A ( T1 and T5 phage resistances)"  
/clone\_lib="NIH\_MGC\_179"  
/notes="Organ: brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV  
(destroyed); Site 2: NotI; Library is oligo-dT primed and  
directionally cloned (EcoRV site is destroyed upon  
cloning). Average insert size 1.1 Kb. Library was  
constructed by (Invitrogen). Note: this is a NIH\_MGC  
Library."

ORIGIN  
Alignment Scores: 3.53e-145 Length: 850  
Pred. No.: 1427.00 Matches: 280  
Score:

Percent Similarity: 98.94% Conservative: 0  
Best Local Similarity: 98.94% Mismatches: 2  
Query Match: 16.94% Indels: 1  
DB: 14 Gaps: 0

US-10-090-458-5 (1-1642) x CD250953 (1-850)

```

Qy 1336 GlyAlaGlyLysSerThrIleLeuAsn-IleLeuValGlyAspIleGluProThrSerGI 1355
Db 3 GGTCTGGCAAAAGCAAAATTTAATANTTCTGGTTGGTATATGAACAACACTTCAGG 62
Qy 1355 YGlnValPheLeuGlyAspTyrsSerGluThrSerGluAspAspSerLeuLysCys 1375
Db 63 CCAGGTATTTTAGGAGATTATCTTCAGAGACAAGTGAAGATGATGATTCATCTGAAGTG 122
Qy 1375 sMetGlyTyrsCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPh 1395
Db 123 TATGGTTACTGTCTCAGATAAACCTTTGTGGCCAGATACATCTGAGGAGAACATTT 182
Qy 1395 eGluIleTyrsGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerAr 1415
Db 183 TGAATTTATGGAGCTCTCAAGGAATGAGTGAAGTGACATGAAGAAGTCATAAGTCG 242
Qy 1415 gIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAl 1435
Db 243 AATAACACATGCATCTGATTTAAAGAACATCTTCAGAGACTGTAAAGAACTACCTGC 302
Qy 1435 aGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyValAsnProGlnIleThrLe 1455
Db 303 AGGAATCAACGAAGTTGTGTTTGTCTAAGTATGCTAGGGAATCTCAGATTACTTT 362
Qy 1455 uLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIle 1475
Db 363 GCTAGATGAACCATCTACAGGTATGATCCAAAGCCAAACAGCACATGTGGCAGCAAT 422
Qy 1475 eArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeuThrThrHisTyrsMetGluGl 1495
Db 423 TCGAACTGCATTTAAACAGAAAGCGGCTGCTATTCTGACCACTCCTATATATGGAGGA 482
Qy 1495 uAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGl 1515
Db 483 GGCAGAGGCTCTCTGTGATCGAGTAGCTATCATGGTGTCTGGGCAGTAAAGATGATCGG 542
Qy 1515 YThrValGlnHisLysSerLysPheGlyLysGlyTyrsPheLeuLysLeuLysLeuLys 1535
Db 543 AACAGTACACATCTAAAGAGTAAATTTGGAAGAGGCTACTTTTGGAAATTAATTTGAA 602
Qy 1535 sAspTrpIleGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrsIlePhePr 1555
Db 603 GGACTGGATAGAAACCTAGAGTAGACCGCTTCAAAGAGAGAAATTCAGTATATTTTCCC 662
Qy 1555 oAsnAlaSerArgGlnGlnSerPheSerSerIleLeuAlaTyrsLysIleProLysGluAs 1575
Db 663 AAATGCAAGCGCTCAGGAAGTTTTTCTCTATTTTGGCTTATATAAATTCCTTAAGAAGA 722
Qy 1575 pValGlnSerLeuSerGlnSerPhePheLeuGluGluAlaLysHisAlaPheAlaIle 1595
Db 723 TGTTCAGTCCCTTTCACAATCTTTTTTAAGCTGGAGAGAGCTAAACATGCTTTTGGCAT 782
Qy 1595 eGluGlyTyrsPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGl 1615
Db 783 TGAAGATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCCTACTAAGA 842
Qy 1615 uGlnGlu 1617
Db 843 ACAAGAG 849

```

RESULT 14  
AY405472 6606 bp DNA linear GSS 16-DEC-2003  
LOCUS AY405472  
DEFINITION Pan troglodytes ABCA1 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
ACCESSION AY405472

Qy 925 IleSerAsp-----LeuIleSerPhePheThrSerGlnAen 936  
 Db 4333 ATTTGGATTATCTGGTGAAGACGTATGTGCAGATCATAGCCAAAGCTTAAAGAACAAAG 4392  
 Qy 937 IleMetValThr----- 940  
 Db 4393 ATCTGGGTGATGAGTTAGGTATGGGGCTTTCCCTGGGTGTCAGTAATACTCAAGCA 4452  
 Qy 941 -----MetIleAenAspSer-----AspTyrValSerVal 950  
 Db 4453 CTTCCTCCGAGTCAAGAAAGTAAATGATCCATCAAAACAAATGAAGAAACACCTAAAGCTG 4512  
 Qy 951 AlaproHisSerAlaAlaLeuAenValMetHisSerGluLysAspTyrVal----- 967  
 Db 4513 GCCAAGACAGTCTCGAGATCGATTTCTCAACAGCTTGGGAAGATTATGACAGACTG 4572  
 Qy 968 -----PheAlaAlaValPheAenSerThrMetValTyrSerLeuProfile 982  
 Db 4573 GACACCAAAATAATGTCAAGGTGTGGTTCAATAACAAGGGCTGGCATCAATCACTCT 4632  
 Qy 983 LeuValAenIleIleSerAsnTyrTyrLeu----- 992  
 Db 4633 TTCTGTAATGTCATCAACATGCCATCTCCGGGCCAACCTCGCAAAAGGAGAGACCTCT 4692  
 Qy 993 -----TyrHisLeuAenValThrGluThrIleGlnIle 1003  
 Db 4693 AGCCATTATGGAATTACTGCTTTCAATCATCCCTCGAATCTCAACCAAGCAG---CAGCTC 4749  
 Qy 1004 TrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGluLeuTyrPhe 1023  
 Db 4750 TCAGAGTGGCTGTGATGACCAACATCAGTGTGATGCTGTGTCATCTGTGTCATCTTT 4809  
 Qy 1024 GlnAlaAlaLeuGluIleValThrAlaMetProTyrPheAlaMetGluAen 1043  
 Db 4810 -----GCAATGCTCTTCCAGCAGCTTGTGCTGATTCCTG 4848  
 Qy 1044 AlaGluAenHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSer 1063  
 Db 4849 ATCCAGAGCGGTCACAAAGCAACACACTGCAGTTCATCAGTGGAGTGAAGCTGTC 4908  
 Qy 1064 AlaTyrTrpIleGlyGlnAlaValAlaValAspIleProLeuPhePheIleIleLeu 1083  
 Db 4909 ATCTACTGGCTCTTAATTTGTCTGGATATGTCAATACCTTGTCTGCTGCCACACTG 4968  
 Qy 1084 MetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeu 1103  
 Db 4969 GTCAATATCATCTTCCTGCTCCAGCAGAGTCCATATGTCCTCCACCAATCTGCCT 5028  
 Qy 1104 AlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAla 1123  
 Db 5029 GTGCTAGCCCTCTACTTTGCTGATGGGTGGTCAATCACACTCTCATGTACCCAGCC 5088  
 Qy 1124 SerPheThrPheLysLysIleLeuAenThrLysGluPheTrpSerPheIleTyrSerVal 1143  
 Db 5089 TCCTTTGTGTCAAG----- 5106  
 Qy 1144 AlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr-----Thr 1161  
 Db 5107 CCCAGCAGCAGCTATGTGGTCTCACCGGTGAGACCTCTTCAATGGCATTAATGGCAGC 5166  
 Qy 1162 IleAlaThr----- 1165  
 Db 5167 GTGGCCACCTTTGTGCTGGAGCTGTTCCCGCAACAATAAGCTGAATAATCAATGATATC 5226  
 Qy 1166 LeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLeuIleSer 1185  
 Db 5227 CTGAAGTCCGTGTC---TTGAATCTCCACATTTTTCCTGGACGAGGGCTCATCGAC 5283  
 Qy 1186 PheIleLys----- 1188  
 Db 5284 ATGCTGAAGAAACCCAGCAATGGCTGTATGCCCTGGAAAGGTTTGGGAGATCCTTTGTG 5343

Qy 1189 -----IleSerTrpLysAenValArgLysAenValAspThrTyrAenProTrpAspArg 1206  
 Db 5344 TCACCATTAICTTGGGACTTGGTGGAGCAACCTCTCTGCC----- 5385  
 Qy 1207 LeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGln 1226  
 Db 5386 ATGCCCTGGAAGGGTGGTGTCTCTCTCATTAATCTGTT-----CTGATCCAG 5433  
 Qy 1227 TyrTyrGluLysLysTyrGlyArgSerIleArgLysAspProPhePheArgAenLeu 1246  
 Db 5434 TAC-----AGATTCTTC-----ATC 5448  
 Qy 1247 SerThrLysSerLysAenArgLysLeuProGluProProAspAenGluAspGluAspGlu 1266  
 Db 5449 AGGCCAGACCTCTAAATGCAAGCTA-----TCTCTCTGAAT---GATGAAGATGAA 5499  
 Qy 1267 AspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGlu 1286  
 Db 5500 GATGTAGCGCGGAACACAGAGATTTCTTGATGGTGGAGGC-----CAG 5544  
 Qy 1287 LysProSerIleMetValSerAenLeuHisGlyGluTyrAspAspLysLysAspPheLeu 1306  
 Db 5545 AATGACATCTTAGAAATCAAGAGCTTGACGAAGATATAT----- 5583  
 Qy 1307 LeuSerArgLysValLysValAlaThrLysTyrIleSerPheCysValLysLysGly 1326  
 Db 5584 -----AGAAAGAGCGGAAGCCTGCTGTGACAGATTTGCGTGGGCATCTCTCTGCT 5637  
 Qy 1327 GluIleLeuGlyLeuLeuGlyProAenGlyAlaGlyLysSerThrIleAenIleLeu 1346  
 Db 5638 GAGTCTTTGGCTCTCTGGAGTTAATGGGCTGGAAATCATCAACTTTCAAGATGTTA 5697  
 Qy 1347 ValGlyAspIleGluProThrSerGlnValPheLeuGlyAspTyrSerSerGluThr 1366  
 Db 5698 ACAGGAGATACCATCTGTACCAGAGAGATGCTTCTCT---AACAAANNNTATCTTA 5754  
 Qy 1367 SerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAenProLeuTrp 1386  
 Db 5755 TCMAACATCCATGAATACATCAGAACATGGCTACTGCTCTAGTTGATGCCATCACA 5814  
 Qy 1387 ProAspThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAla 1406  
 Db 5815 GAGCTGTGTGATGGAGAGAACACGTGAGTCTTGTGCTTTGAGAGAGATCCACAGAG 5874  
 Qy 1407 SerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeu 1426  
 Db 5875 AAAGAGTGGCAAGTTGGTGGCGGATTCGGAACCTGGGCTCGTGAAGTATGGA 5934  
 Qy 1427 GlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSer 1446  
 Db 5935 GAAAAATATGTGTGTAATATAGTGGAGGCAACAAACGCAAGCTCTCTACAGCCATCGCT 5994  
 Qy 1447 MetLeuGlyAenProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLys 1466  
 Db 5995 TTGATCGCGGCTCTCTGCTGTGTCTGTGATGAACCCACACAGCATGGATCCCAAA 6054  
 Qy 1467 AlalysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAenArgLysArgAlaAla 1486  
 Db 6055 GCCCGCGGTTCTTGTGAATGTGCCCTAAGTGTGTCAAG---GAGGGGAGATCAGTA 6111  
 Qy 1487 IleLeuThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMet 1506  
 Db 6112 GTGCTTACATCTATAGTATGAAGATGTGAAGCTCTTGTCACTAGATGGCAATCATG 6171  
 Qy 1507 ValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLys 1526  
 Db 6172 GTCAATGAAGGTTTCAAGTGTGGAGTGTCCAGCATCTTAAATAAATAGTTTGGAGAT 6231  
 Qy 1527 GlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAenLeuGluValAspArgLeu 1546  
 Db 6232 GGTATACAATAGTTGTACGAATAGCAGGG-----TCCAAACCCGACCTAGACCTGTC 6285  
 Qy 1547 GlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSerSerIle 1566

```
Db 2176 GCCTTTGGGTTGGCTGTGAGTACTTGGCCCTTTTGGAGGACGAGCGCATGGAGTGCAG 2235
Qy 386 PheSerAsnLeuThrAlaGlyPro-----TyrProLeuIleIleThrIle 400
Db 2236 TGGGACAACTGTTGAGAGTCTCTGGAGGAAGATGGCTTCAATCTCAACCTTCGGTC 2295
Qy 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuAlaValTyrLeuAspGlnVal 420
Db 2296 TCCATGATGCTGTTGACACCTTCCCTATAGGGGTGATGACCTGGTACATGAGGCTGC 2355
Qy 421 IleProGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440
Db 2356 TTTCCAGCCGACGACGAAATCCAGGCCCTGTATTTCTTGGACCAAGTCTACTGG 2415
Qy 441 -----SerLysSerLysArgAsnTyrGluGluLeuSerGlu 452
Db 2416 TTTGGCGAAGAGTGTAGAGAGAGCCACCTGGTTCACCAAGAGAGATATACAGAA 2475
Qy 453 GlyAsnValAsnGlyAsnIleSerPheSerGluIleIleGluProValSerSerGluPhe 472
Db 2476 -----ATCTGCATGGAGGAG-----GAACCCACCCACTTG----- 2505
Qy 473 ValGlySerGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGlu 492
Db 2506 -----AAGCTGGCGGTGTCATTCAGAACCTGTAAAGTCTACCGA---GATGGGATG 2556
Qy 493 AsnValGluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeu 512
Db 2557 AAGGTG---GCTGTGATGGCTGGCACCTGAATTTTATGAGGCCAGATCACTCTCTTC 2613
Qy 513 LeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysPro 532
Db 2614 CTGGCCACAACTGAGCGGGAGAGACGACCACCATGTCAATCTGACCGGGTTGTTCCTCC 2673
Qy 533 ProSerAspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPhe 552
Db 2674 CCGACCTCGGGCACCGCTACATCTCTGGGAAAAGACATTCGC-----TCTGAGATGAGC 2727
Qy 553 GluAlaArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThr 572
Db 2728 ACCATCGGCAGAACCTGGGGGTCTGTCCCCACATAAAGTGTGTGTGACATGTGACT 2787
Qy 573 ValGluGluAsnSerIleLeuAlaSerIleLysGlyIleProAlaAsnAsnIleIle 592
Db 2788 GTCGAGAACACATCTGTTCTATGCCCGCTTGAAGGGCTCTCTGAGAGCAGCTGAAG 2847
Qy 593 GlnGluValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLys---AspAsnGln 611
Db 2848 GCGGAGATGAGCAGATGGCCCTGGATGTGTGTTTGCATCAAGCAAGCTGAAAGCAAA 2907
Qy 612 AlaLysLysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGly 631
Db 2908 ACAAGCCAGCTGTACGTGGAATGCGAGAAAGCTATCTGTGGCCTTGGCCCTTGTGCGG 2967
Qy 632 AsnProLysIleLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHis 651
Db 2968 GGATCTAAGTTCTCATCTGGATGAACCCACAGCTGGTGTGGACCTTACTCCGCGAG 3027
Qy 652 IleValTrpAsn---LeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThr 670
Db 3028 GGAAATGGGAGCTGTCTGCTGAATACCGACAA---GGCCGACACCATTAATCTCTACA 3084
Qy 671 HisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMet 690
Db 3085 CACCACATGATGAAGCGGACGTCCTGGGGGACAGGATTCATCTCCCATGGGAAG 3144
Qy 691 LeuLysCysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrArgLeu 710
Db 3145 CTGTGCTGTGTGGCTCTCCCTGTTTCTGAAGAACCAAGCTGGGAACAGGCTACTACCTG 3204
Qy 711 SerMet----- 712
: : : : :
Db 3205 ACCTTGGTCAAGAAAGATGTGGAATCTCCTCAGTTCTCTGCAGAAAACAGTAGTACACT 3264
Qy 713 -----TyrIleAspLysTyrCysAla----- 719
Db 3265 GTGTCATACCTGAAAGAGGAGGACAGTGTCTCTCAGAGCAGTCTTGATGTGGCTGGGC 3324
Qy 720 -----ThrGluSerLeuSerSerLeuValLys 728
Db 3325 AGCGACCATGAGAGTGACACGCTGACCATCATCTCTCTGTATCTCAACCTCATCAGG 3384
Qy 729 GlnHisIleProGlyAlaThrLeuLeuGlnIleAsnAspGlnGlnLeuValTyrSerLeu 748
Db 3385 AAGCATGTGTCTGAAGCCCGGCTGTGTGAAGACATAGGCGATGAGCTAGCTATGTGCTG 3444
Qy 749 ProPheLysAspMetAspLys-----PheSerGlyLeuPheSerAlaLeuAspSerHis 766
Db 3445 CCATATGAAGCTGCTAAGGAGGAGCGCTTTGTGAACTCTTTTCATGATTTGATGACCGG 3504
Qy 767 ---SerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrIleGluAspValPhe 785
Db 3505 CTCTCAGACCTGGGCATTTCTAGTTATGGCATCTCAGAGACGACCCCTGGGAAGAAATATTC 3564
Qy 786 LeuLysLeuGluValGluAlaGluIleAsp----- 795
Db 3565 CTCAGGTGGCGGAGAGAGAGTGGGTGGATGCTGAGACCTCAGATGCTTACCTGCCAGCA 3624
Qy 796 -----GlnAlaAspTyrSerValPheThrGln--- 804
Db 3625 AGACGAAACAGCGCGGCTTGGGGGACAGCAGAGCTGTCTCGCCGTTTCACTGAAGAT 3684
Qy 805 -----GlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGln 821
Db 3685 GATGCTGCTGATCAATGATTTCTGCATAGACCCAGATCC-----AGAGAGACA 3735
Qy 822 SerLeuLeuIleLeuSerGluThrLysAlaSerLeu---ValSerThrMetSerLeuTrp 840
Db 3736 GACTTGTCTAGTGGATGGATGGCAAGGGTCTTACCAGGTGAAGGCTGGAACCTTACA 3795
Qy 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860
Db 3796 CACAACAGTTTGGGCCCTTTTGTGAAGAGACTGCTAATGCCACGAGGATCGGAAA 3855
Qy 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880
Db 3856 GGATTTTGTCTCAGATTGTCTCCAGCTGTGTTTCTCTGCATGTGCCCTTGTGTTTCCAG 3915
Qy 881 LeuVal----- 882
Db 3916 TGTATGCTGGCACCTTTGGCAAGTACCCAGCCTGGAATTTCAGCCCTGGATGTACAAC 3975
Qy 882 ----- 882
Db 3976 GAACAGTACATTTGTTCAGCAATGATGCTCCTGAGGACACGGGAACCTTGGAACTCTTA 4035
Qy 882 ----- 882
Db 4036 AACGCCCTCACCAGAACCCCTGGCTCGGACCCGCTGTATGGAGGAAAACCCCAATCCCA 4095
Qy 883 -----HisHisSerPheLysAsnAlaValProIleLysLeu 895
Db 4096 CACAGCCCTCCAGGCGGGGAGAGAGTGGNACCACTGCCCAGTTCCC---CAGACC 4152
Qy 896 ValProAspLeuTyr----- 900
Db 4153 ATCATGACCTCTTCCAGNATGGGAATGGACAATGCAACCCCTTCACTGTCATGCCAG 4212
Qy 901 -----PheLeuLysProGly----- 905
Db 4213 TGTAGCGGACAAAATCAAGAGATGCTGCTGTGTGTCTCCCGGAGGGGGGGCTG 4272
Qy 906 ---AspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAspSerAsp 924
Db 4273 CCTCCTCCACAAAGAAAACAAACACTGTCAGATATCTTTCAGGACCTTGACAGGAAGAAC 4332
```

Db 5632 CCGGAGAGTCTTTGGACTCTGGGAGTTAAGGAGCTGGGAAGTCAACAACTTCAAG 5691  
 Qy 1345 IleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSer 1364  
 Db 5692 ATGCTGACTGAGACACCCCTGTGACGAGGAGATGCTTCAACACATCTTATCA 5751  
 Qy 1365 GluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAspPro 1384  
 Db 5752 AATATCATGAAGTACACCAAGAC-----ATGGCTACTGCCCTCAGTTGACGCC 5802  
 Qy 1385 LeuTrpProAspThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMet 1404  
 Db 5803 ATCACAGAGCTGCTGAGGAGAGCATGTGGAGTTCTTCCCTCCCTGAGGGAGTC 5862  
 Qy 1405 SerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGlu 1424  
 Db 5863 CCAGAAAGGAAGTTGGCAAGTTGGTGAATGGCAATTCGCAACTGGGCCCTGGTAAAG 5922  
 Qy 1425 HisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuLysPheAla 1444  
 Db 5923 TATCGAGAAATAATATGCCAGTAATACAGTGGCGGCAACAAAGAAAGCTCTCCACAGCC 5982  
 Qy 1445 LeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAsp 1464  
 Db 5983 ATGGCTTTGATGGCGGACCTCCCTGTGTGTCTGATGATGAACCAACACAGCATGGAC 6042  
 Qy 1465 ProLysAlaLysGlnHisMetTyrPargAlaIleArgThrAlaPheLysAsnArgLysArg 1484  
 Db 6043 CCTAAGCCCGAGATTCTTGTGAATTTGCTGAATTTGCTGAATTTGCTGAATTTGCTGA 6099  
 Qy 1485 AlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAla 1504  
 Db 6100 TCTGTAGTCTTATCATCTCATAGTATGGAGATGTGAAGCTCTTGTACAGGATGGCC 6159  
 Qy 1505 IleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPhe 1524  
 Db 6160 ATAATGTGTCATATGGAAGTTTCAGTGGCTTGGCAGTGTCCAACTCTGAAACACAGGTTT 6219  
 Qy 1525 GlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeuGluValAsp 1544  
 Db 6220 GGAGATGGTTATACATAGTTGTAGATAGCAGC-----TCCACCTGACCTGAAG 6273  
 Qy 1545 ArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSerPheSer 1564  
 Db 6274 CTTGCCAGGAGTCTTTGGACTTTCGCTTTCGGGAAAGTCTCTTAAAGAAACATCGA 6333  
 Qy 1565 SerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePhe 1584  
 Db 6334 AACATGCTTCAGTACCACTTCCA---TCCCTCTGTCACTCTAGCAGGATATTGAC 6390  
 Qy 1585 LysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThr 1604  
 Db 6391 ATCCCTCTCCAGAGCAAAAGCGACTCCACATAGAGACTACTCTGTCTCTCAGACAA 6450  
 Qy 1605 LeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAspAsn----- 1621  
 Db 6451 CTGTACCAAGTATTGTGAATTTGCCAAGGACCAAGTGAATGATGACCTTAAAGAC 6510  
 Qy 1622 -----SerCysGlyThrLeuAsnSerThrLeuTrp 1631  
 Db 6511 CTGTCACTGCACAAAACACAGACAGTTGTGGATGGCGGCTTCTCATCTCTTTTGCAG 6570  
 Qy 1632 TrpGluArgThrGlnGluAspArgVal 1640  
 Db 6571 GATGAGAAAGTGAAGAAAGTTATGTA 6597

RESULT 12  
 AY405471  
 LOCUS Homo sapiens ABCA1 gene, 6606 bp DNA linear GSS 16-DEC-2003  
 DEFINITION Homo sapiens ABCA1 gene, partial sequence, genomic survey sequence.  
 ACCESSION AY405471  
 VERSION AY405471.1 GI:39761445

KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (Bases 1 to 6606)  
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE 2 (Bases 1 to 6606)  
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
 COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.  
 FEATURES  
 Location/Qualifiers  
 1..6606  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>6606  
 /gene="ABCA1"  
 /locus\_tag="HQM2235"  
 Source  
 gene  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,42e-145 Length: 6606  
 Score: 1444.00 Matches: 450  
 Percent Similarity: 43.41% Conservative: 291  
 Best Local Similarity: 26.36% Mismatches: 556  
 Query Match: 17.14% Indels: 410  
 DB: 29 Gaps: 55  
 US-10-090-458-5 (1-1642) x AY405471 (1-6606)  
 Qy 240 IleValAlaGluLysGluLysLysLysLysGluPheLeuLysIleMetGlyLeuHisasp 259  
 Db 1795 ATCGTGTATGAGAGGAGCAGCGCTGAAGAGACCATCGGATCATGGGCTGCAAC 1854  
 Qy 260 ThrAlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMet 277  
 Db 1855 AGCATCCTCTGTTAGTGTGTTCAIT-----AGTAGCCTCATCTCTCTTGTGAGC 1908  
 Qy 278 ---SerLeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSer 296  
 Db 1909 GCTGGCTGCTAGTGTCTCATCTGAAAGTAGGAAACCTGCTG---CCCTACAGTATCCC 1965  
 Qy 297 IleValIlePheLeuPhePheLeuTyrGlyLeuSerValPhePheAlaLeuMet 316  
 Db 1966 ACGGTGGTGTGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2025  
 Qy 317 LeuThrProLeuPheLysSerLysHisVal-----GlyIleValGluPhe 332  
 Db 2026 ATTACACACTCTTCTCCAGACCAACCTGGCAGCAGCTGTGGGGGATCATCTACTTC 2085  
 Qy 333 PheValThrValAlaPhe-----GlyPheIleGlyLeuMetIle 345  
 Db 2086 ACGCTGTACCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2145  
 Qy 346 IleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCys 365  
 Db 2146 AGATCTTCGCTAGC-----CTGCTGCTCTCT-----GTG 2175  
 Qy 366 ThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSer 385

QY 805 -----GlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 3682 GATGATCTGTGTATCCCAATGATCTGACATAGACCCAGATCC-----AGGAG 3732  
QY 821 GlnSerLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
Db 3733 ACCGACCTGTCAGTGGGATGGAGCGGCAAGGCTCTAC-----CAGCTGAAGGCTGG 3786  
QY 841 Lys-----GlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGlu 858  
Db 3787 AAATCACCACCAACAGATTTGTGGCCCTTTTGTGGAGAGGCTGTGTGATTCGACAGCG 3846  
QY 859 SerLysSer-----ValArgSerValLeuLeuLeuLeuLeu 871  
Db 3847 AGCCGAGAGGTTTCTTGTCTCAGATTGCTGCTGCCAGCTGTCTTGTTCATGCGCTG 3906  
QY 872 PhePheThr-----ValGlnIlePhe 878  
Db 3907 GTCTTCAGCTGATGTGCCACCTTTGGCAAGTACCCAGCTGGAACCTTACGCTGG 3966  
QY 879 MetPheLeuValHisHisSerPheLysAsnAlaValValProIleLysLeu----- 895  
Db 3967 ATGTATATGAGCAGTATACATTTGTGATTAATGATGCTCCGAGGACATGGGCACCCAG 4026  
QY 895 ----- 895  
Db 4027 GAATCCTGAATGCTCTGACCAAGATCCAGGCTTTGGGACCCGCTGTATGGAAGAAAC 4086  
QY 896 ---ValProasp----- 898  
Db 4087 CCAATCCAGATACCCCTTGTGCTGGTGGGAGGAGTGGACCATCAGCCCGCTCCC 4146  
QY 898 ----- 898  
Db 4147 CAGAGCATCGTGACCTTCCAGATGGAACCTGGACCATGAGACCCCTCACCTGG 4206  
QY 899 -----LeuTyrPheLeuLysProGly----- 905  
Db 4207 TGCAGTGTAGCAGTGACAAATCAAGAAATGCTGCTGTGTGTCCTCCGAGCGGCGAGG 4266  
QY 906 -----AspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAsp 922  
Db 4267 GGGCTGCCACTCTCCAGAGAAACAGAAACCGCAGACATCTTCAGATCTGACAGGA 4326  
QY 923 SerAspIleSerAsp-----LeuIleSerPhePheThrSer 934  
Db 4327 AGAAATATTTTCAGATTAACCTGCTGAGACGTCAGTGCAGATCATAGCAAGAGCTTAAAG 4386  
QY 935 GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
Db 4387 AATAAGATCTGGTGAACGAGTTTCGGTATGGCGGGTTTCCCTGGGTGTCAGTAATTCT 4446  
QY 955 AlaAlaLeuAsnValMetHisSerGluLysAspTyrVal----- 957  
Db 4447 CAAGCACTTCTCCGAGCCATGAAGTTAATGATGCTATCAAGCAAAATGAAGAACTCCTG 4506  
QY 968 -----PheAlaAla 970  
Db 4507 AAGCTGACCAAGACACCTCGCAGATCGCTTCTCAGACGCTGGGAAGGTTCTAGGCA 4566  
QY 971 -----ValPheAsnSerThrMetValTyrSerLeu 980  
Db 4567 GGGCTGGATACGAAAAACAATGTCAAGGTGTGGTTCAATAACAAGGCTGGCATGCTATC 4626  
QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeu----- 992  
Db 4627 AGTTGCTCTGAATGTCATCAACATGCCATTCTCCGGGCCAACCTGCGAAGAGGAGAG 4686  
QY 993 -----TyrHisLeuAsnValThrGluThrIle 1001  
Db 4687 AACCCGAGCCAGTATGGAATCACTGCTTTCAACACCCCTTTGAACCTCACTAATACAG--- 4743  
QY 1002 GlnIleTrpSerThrProPhePheGlnIleThrAspIleValPheLysIleGluLeu 1021

Db 4744 CAGCTCTCAGAGGTGGCTCTGATGACACCTCTGTGACGCTCTGTCTATCTGTGTC 4803  
QY 1022 TyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAlaMet 1041  
Db 4804 ATCTTT-----CGATGTCTTGTCTCCCTGCCAGCTTTGTTGTG 4842  
QY 1042 GluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeu 1061  
Db 4843 TTCTGATCCAGGAGCGTGTGAGCAAGCCAGCATCTTCAGTTCATCAGCGGCTGAAG 4902  
QY 1062 ProSerAlaTyrTyrIleGlyGlnAlaValValAspIleProLeuPhePheIleLeuLeu 1081  
Db 4903 CCTGTCACTACTGCTGCTCCAAATTTCTGCGGATATGTCAATATGTGCTCCCTGCT 4962  
QY 1082 IleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVallys 1101  
Db 4963 ACATCGGTCAATTATCATCTTCATCTGCTCAGCAGAAAGTCTATGTCTCTACCAAC 5022  
QY 1102 PheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyr 1121  
Db 5023 CTGCCCCGTCTAGCCCTTCTGCTCTGTGTATGGTGTGATGATCCTCTCATGTAT 5082  
QY 1122 IleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyr 1141  
Db 5083 CCAGCATCTTTGTGTTCAAG----- 5103  
QY 1142 SerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr--- 1160  
Db 5104 ---ATCCCCAGACCGCTATGTGTTCTCACCAGTGAACCTCTTCATCGCATCAAT 5160  
QY 1161 ---ThrIleAlaThr----- 1164  
Db 5161 GGCAGTGTGGCCACTTTCGTACTGGAGCTGTTTACAAACAATAAGCTCAATGACATCAAT 5220  
QY 1165 ---IleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLeu 1183  
Db 5221 GACATCTCGAGTCTGTGTT---CTTATCTCCACATTTTTCGCTGGGCGAGGGCTC 5277  
QY 1184 IleSerPheIleLys----- 1188  
Db 5278 ATCGACATGGTGAGAACACGAGCCGATGCGCTGAGAGGTTTGGGAGAACCGC 5337  
QY 1189 -----IleSerTrpLysValArgLysAsnValAspThrTyrAsnProTrp 1204  
Db 5338 TTCTGCTCTCCGCTCTCTTGGACTTGTAGCAGCGAACCTT----- 5379  
QY 1205 AspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTyrIlePheLeu 1224  
Db 5380 -----TTTGCCATGCGCGTG-----GAAAGGGTGGTCTTCTCTCAT 5418  
QY 1225 LeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAspProPhePheArg 1244  
Db 5419 ACTGTTCTGATCCAGTACAGATTT-----TTTCATCAGG 5451  
QY 1245 AsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAspAsnGluAspGlu 1264  
Db 5452 CCCACACCTGTAAAGCGAG-----CTTCTCTCTTCAAT---GACGAG 5493  
QY 1265 AspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCys 1284  
Db 5494 GATGAGGATGTGAGCGGAGAGACAGAGGATCTGGACCGGTGGCGGA----- 5541  
QY 1285 GluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAsp 1304  
Db 5542 ---CAGAATGCATCTAGATCAAGAACTGACCAAGATCTAT----- 5583  
QY 1305 PheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCysVallys 1324  
Db 5584 -----AGGAGGAACGGAAGCCTGAGTCGACAGATCTGATCGCATCCCT 5631  
QY 1325 LysGlyGluIleLeuGlyLeuGlyProAsnGlyAlaGlyLysSerThrIleLeuAsn 1344

Thu Apr 15 09:20:54 2004

Alignment Scores:

Alignment Scores:			
Pred. No.:	1.54e-148	Length:	6600
Score:	1471.00	Matches:	458
Percent Similarity:	43.26%	Conservative:	290
Best Local Similarity:	26.4%	Mismatches:	559
Query Match:	17.4%	Indels:	422
DB:	29	Gaps:	57

US-10-090-458-5 (1-1642) x AY405473 (1-6600)

Qy	223	IleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIle-----His	239
Db	1744	CTCTTCATGCTCTAGCCTGGATC-----TACTCTGCTGTCATCATCAAGAGC	1794
Qy	240	IleValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAsp	259
Db	1795	ATTGTGTATGAGAAGAGGCTGGCTGAAGAGACCATGGGATCATGGTCTGGACAAT	1854
Qy	260	ThrAlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMet----	277
Db	1855	GGCATCTCTGGTTTAGTGGTTGGTT-----AGCAGCCTCATCCCTCTGCTGTGTGAGC	1908
Qy	278	---SerLeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSer	296
Db	1909	GCTGGCTGCTGGTGGTCACTCTGAAGTAGGAACCTGCTG-----CCCTATAGTACCCC	1965
Qy	297	IleValIlePheLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMet	316
Db	1966	AGCGTGGTGTTCGCTTCCTGCTGCTGTTGGCCATGCTGACCATCTCTACAGTGTCTCCTC	2025
Qy	317	LeuThrProLeuPheLysSerLysHisVal-----GlyIleValGluPhe	332
Db	2026	ATTAGCAGCTCTCTCTCCGTGGCAACCTGGCACACCGTGTGGGGGCACTCACTACTTC	2085
Qy	333	PheValThrValAlaPhe-----GlyPheIleGlyLeuMetIle	345
Db	2086	ACGCTGTACCTGCCCTATGCTGCTGCTAGCCTGGCAGGACTATGTGGGCTTCTCCATC	2145
Qy	346	IleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCys	365
Db	2146	AAGACTTTTGTGTAGC-----CTGTGCTCTCTT-----GTG	2175
Qy	366	ThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSer	385
Db	2176	GCTTTTGGATTCCGCTGTGAGTATTTCGCCCTTTTCAGGAGCAAGGTATCGGGGTCAA	2235
Qy	386	PheSerAsnLeuThrAlaGlyPro-----TyrProLeuIleIleThrIle	400
Db	2236	TGGACAACTCTTTTGAGAGCCCGTGGAGGAGGAGCGGCTTCAATCTCACCATCTGCAGTG	2295
Qy	401	IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal	420
Db	2296	TCCATGATGCTCTTTGACACCTTCTCTATGGCTGTGATGACATGGTACATCAAGACCGCTC	2355
Qy	421	IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp	440
Db	2356	TTTCCAGACAGTAGTAGTAATTCACAGCCCTGGTATTTCCTGTGACCAAGTCACTATCGG	2415
Qy	441	SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyLeuAsnValAsnGlyAsnIle---	459
Db	2416	TTTGGT-----GAGGAATTCATGAGAAGACCCCTCTGGTTCAGCCAG	2460
Qy	460	---SerPheSerGluIleIle-----GluProValSerSerGluPheValGlyLys	475
Db	2461	AAGGGAGTGTGAGAAATCTGCATGGAAGAGAACCCACTCATCTGAGCGCTG-----	2511
Qy	476	GluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGlu	495
Db	2512	---GGGGTGTCCATTCAGAACCTGCTGAAGCTTTACCGA---GATGGCATGAAGGTT---	2562
Qy	496	AlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHis	515
Db	2563	GCTGTGATGGCTTGGCCCTCAACTTTTACCAAGCCAGATTACTCTCTCTCTGGCCAC	2622

QY	516	SerGlyThrGlyIysSerThrLeuMetAsnIleLeuCysGlyValLeuCyseProProSerAsp	535
DB	2623	AAATGGACGAGGGAGAACACACACCACTGTCAAATACGTACCTGGGCTGTTCCTCCCACTTCT	2682
QY	536	GlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArg	555
DB	2683	GGCACGGCTACATCCTGGGGAAGGACATTCGCGAGATGAGCTCCATCCCG	2736
QY	556	LysMetIleGlyIleCysProGlnIleuAspIleHisPheAspValLeuThrValGluGlu	575
DB	2737	CAGAACTGGGAGTCTGCCACGATAATATGCTGTTTGACATGCTGACTGTGCGAAGAG	2796
QY	576	AsnLeuSerIleLeuAlaSerIleTyrGlyIleProAlaAsnAsnIleIleGlnGluVal	595
DB	2797	CACATCTGGTTCATCGCGCCTAAAGGGGCTCTCAGAAAGCAGCCTGAAAGCAGAGATG	2856
QY	596	GlnIysValLeuLeuAspLeuAspMetGlnThrIleLys--AspAsnGlnAlaLysLys	614
DB	2857	GAGCAGATGCCCTGGATGTGGCTTACCCCGAGCAGCTGNAAGCAAAACGAGTCAG	2916
QY	615	LeuSerGlyGlyGlnIysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLys	634
DB	2917	CTCTCAGTGGGATGCAGAGAAAGCTGTCTGGCCCTTGGCCTTCTGTGGGTGATCCAAG	2976
QY	635	IleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleValTyr	654
DB	2977	GTTCATCTTGCAGAGCCCAACGCGGGTGGACCCGTACTCTGCAGGGGAAATATGG	3036
QY	655	Asn---LeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPheMet	673
DB	3037	GAATCCTCTGCTAAATACCGCAA---GGCGCGACCATATTATTTGTCTACACACCACATG	3093
QY	674	AspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLysCys	693
DB	3094	GACGAAGCTGACATCTCTGGGGACAGATTCGCTCATTTCCATGGGAAGCTGCTTGT	3153
QY	694	ValGlySerSerMetPheLeuLysSerLysTyrGlyIleGlyTyrArgIleuSerMetTyr	713
DB	3154	GTGGCTCTCTCCCTGTTTTGAAAAAACCAAGTTGGGAACGGGTTACTATCTGCACCCCTGGT	3213
QY	713	-----	713
DB	3214	AAGAAAGATGTGAATCGTCCCTCAGTTCTCTGCAGAAACAGTAGCAGCACCGTGTCTGT	3273
QY	713	-----	713
DB	3274	CTGAAAAAGGAGGACAGTGTTCACAGCAGATTCTGATGCTGGCTGGCGACGACCAT	3333
QY	714	-----IleAspLysTyrCysAlaThrGluSerLeuSerSerLeuVal	727
DB	3334	GAAGTGACACGTGACCATTCGAT-----GTCTCTGTATCTCCACCTCATC	3381
QY	728	LysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnIleuValTyrSer	747
DB	3382	AGGAAGCACGTGTCTGAAGCCCGCTGGTGAGGACATTTGGGACAGCAGCTGACCTATGT	3441
QY	748	LeuProPheLysAspMetAspLys-----PheSerGlyLeuPheSerAlaLeuAspSer	765
DB	3442	CTGCGGTACGAAGCCGGAAGGAGGAGCCCTTTGTGGAATCTTTCATGAGATGTGATGAC	3501
QY	766	His---SerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspVal	784
DB	3502	CGGCTCTCAGACCTGGGCATCTCCAGTATTGGCATCTCGGAGAGCCACCTCGGAAGAATA	3561
QY	785	PheLeuLysLeuGluValGluAlaGluIleAsp-----	795
DB	3562	TTCTCTCAAGTGGCTGAAGAGAGCGGGTGGATGCTGAGACCTCAGATGGTACTTTCGCCA	3621
QY	796	-----GlnAlaAspTyrSerValPheThrGln	804
DB	3622	GCAAGACGAACACAGCGGGCTTCGGGGACAACAGCAGAGCTGTCTGCACCATTTACGGAA	3681





```
QY 429 gSerSerLeuTyrPheLeuLysProSerTyrTrpSerLysSerLysArgAsnTyrGluG1 449
Db 1453 GTCTCCATGGTTTTCTCGAAGTGGTCTTACTGTTTAAACCGACAGAGGCTAGTCACTG 1512
QY 449 uLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGluLeuIleGluProValSe 469
Db 1513 GGTCTTGAGAAATAGATTCGGATCCCTCGCTGAATGACTCACTGGAACACGAGTGC 1572
QY 469 rSerGluPheValGlyLysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLy 489
Db 1573 TCCAGAAATTCAGGGGAAGAGCCATCAGATCAAAATCTTAAAGAAATATTCTCG 1632
QY 489 sLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnI1 509
Db 1633 AAAGCATGGGAAGTAGAAGCTTTGAGAGGCTCTGGGTTTGACATATATGAAGCCAGAT 1692
QY 509 eThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysG1 529
Db 1693 CACTGCCCTCTCTGGTCACTGAGCTGGGGAACACACCTGCTGATAATACACTCACTGG 1752
QY 529 yLeuCysProSerAspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAs 549
Db 1753 ACTGTCAACCCCTACTCAGGTTCCGTCAACATTTATAACACAGACAGTTTCAGAGATGA 1812
QY 549 pGluMetPheGluAlaArgLysMetIleGlyLeCysProGlnLeuAspIleHisPheAs 569
Db 1813 CGACTCTGACCGCTCTCAGATCAACGGAGTTTGCCACAGTCCACAGTCCAGTACAGTTGG 1872
QY 569 pValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAs 589
Db 1873 CTTTCTCACTCCGAGAAAACCTCAGGCTTTTGCCCAAGATAAAGGAATTTGCCACA 1932
QY 589 nAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLysAs 609
Db 1933 TGAAGTGAACAAGAGGTACACAGATTCTACAGACTTAGAAATGGAATAATTCAAGA 1992
QY 609 pAsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaVa 629
Db 1993 CATTTCTGCTCAGAAATTTAAGTGGCGGACCAAAAGGAAGTTGACTCTTGGGATCGCAT 2052
QY 629 lLeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSe 649
Db 2053 TTTAGAGACCCCGAGGTTTTCCTCGTGGATGAACCAACTCTGGGTGGATCCACTCTC 2112
QY 649 rArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSe 669
Db 2113 GAGGCATCGCATATGGAACCTACTGAAAGAGAGAGAGAGAGAGAGAGTAAATGTCTTCAG 2172
QY 669 rThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnG1 689
Db 2173 TACTCAGTTCATGGATGAGGCTGATATCATGCCGACAGGAAGGTGTTTATATCCAAATGG 2232
QY 689 yMetLeuLysCysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrAr 709
Db 2233 GAGACTGAAGTCTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2292
QY 709 gLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysG1 729
Db 2293 TCTAAGTTTGAATTTGAATGAAGCGTGTGATCCAGAGGGTATACAGTCACTGGTCAAAA 2352
QY 729 nHisIle-ProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeup 749
Db 2353 GCACATCTCCGACCGCGCTGACCGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2412
QY 749 roPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnL 769
Db 2413 CCCTGGAAAGGACAAACAAGTTTCCAGACCTTTACCGGATCTTTGATAGATGTTCTCAACC 2472
QY 769 euGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLysLeuG 789
Db 2473 AAGGCATTGAGGATATCGCGTTTCCATGACACCTGACAACTGAATGAGGTGTTCTGAAGTAG 2532
```

```
QY 789 luValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluG 809
Db 2533 AAGGGAAGTCGATGGCTGATGATCAGATTTGGAAATTTGGGGCGCTTCAAAAGCGATG 2592
QY 809 luGluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuLeuLeuSerGluT 829
Db 2593 GGGCAAGACACATGGAAGACCTCTGTTGAGCTGGAGCAAGTGTGTCTTTA-----GATT 2646
QY 829 hrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGln---MetTyrThrIleA 848
Db 2647 CGTCAGGAAGCTCAGTGAAGTGGCTGTGGAGGACAGAGCTCTGGCGAGCCG-G 2705
QY 848 laLysPheHisPhePheThrLeuLysArgGluSerLysSerValArgSerValLeuLeuL 868
Db 2706 CGAAGTTCGCTCTTCAAGCTCAAGAATGAACGGAAGAGTCTGATGATGTGTATTTC 2765
QY 868 euLeuLeuIlePhePheThrValGlnIlePheMetPheLeuVal-----HisH 884
Db 2766 TTTTGGGCATTAGCTTTGTCCCCCAACTTTTGGAAACATCTCGTCTACAAGGTATACCACA 2825
QY 884 isSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLysP 904
Db 2826 AAAGTTATTTCATGGGGG-----CTATCTCCAAGTATGTTATTTCTTTCAC 2870
QY 904 roGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerA 924
Db 2871 CAGGACGCCACACAGGATCCTCTGACCCACTTACTGTCATCATATAGAACAGGTTCAA 2930
QY 924 spIleSerAspLeuLysSerPhePheThrSerGlnAsnIleMetValThrMet-----I 942
Db 2931 GATCGATAAAGTTTGTACAGCCTAAGACAGAGGCGATAGCATTAGACGTGATGCC 2990
QY 942 leAsnAspSerAspTyr-ValSerValAlaProHisSerAlaAlaLeuAsnValMetHis 961
Db 2991 TGGGAGCGCAGAAATGGACAGAGGCGGTGTACATGTGTCTATCCCGTCTGGGT 3050
QY 962 SerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeuPro 981
Db 3051 GAAGAAAAGGCTCTCAGATTTTCCGTAGCGTGTATGCTGCAAAAGACTAACTGTCTCCA 3110
QY 982 lIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIle 1001
Db 3111 GTCCTCATGGATATTATTAGCATGGGTCTGTGGAATATTAACTCTCTTGAAGCGATC 3170
QY 1002 GlnIleTrpSerThrProPhePheGlnGlu-----IleThrAspIleValPheLysIle 1019
Db 3171 CAGACCGACAGAGCAGCAGTCTTTCGAGACACACATGCTTTATGATACGGATTATGAGC 3230
QY 1020 GluLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyrPhe 1039
Db 3231 AAGCATTTTCTGGATACCA-----GTGGCTGCAGGCTCACCCCTTACATT 3278
QY 1040 AlaMet 1041
Db 3279 GCGATG 3284
RESULT 10
U66672 Homo sapiens clone EST90625 mRNA sequence.
LOCUS U66672
DEFINITION Homo sapiens clone EST90625 mRNA sequence.
ACCESSION U66672
VERSION U66672.1 GI:1906557
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1346)
AUTHORS Allikmets,R., Gerrard,B., Hutchinson,A. and Dean,M.
TITLE Characterization of the human ABC superfamily: isolation and
mapping of 21 new genes using the expressed sequence tags database
JOURNAL Hum. Mol. Genet. 5 (10), 1649-1655 (1996)
MEDLINE 97049974
```

```

REFERENCE 6 (bases 1 to 3300)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayaehida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration and Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gscl.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]
COMMENT cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
FEATURES
source 1..3300
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:D630040K07"
/db_xref="MGI:2422672"
/db_xref="taxon:10090"
/clone="D630040K07"
/tissue_type="kidney"
/clone_libs="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
misc_feature 158..53299
/note="putative
similar to CDNA FLJ32506 FIS, CLONE SMINT1000042, WEAKLY
SIMILAR TO ATP-BINDING CASSETTE, SUB-FAMILY A, MEMBER 3
(FRAGMENT) [Homo sapiens] (SPTR|O96MD8, evidence: FASTY,
77.5%ID, 100%length, match=2568)"
ORIGIN
Alignment Scores:
Pred. NO.: 6.4e-173 Length: 3300
Score: 1688.50 Matches: 409
Percent Similarity: 57.14% Conservative: 199
Best Local Similarity: 38.44% Mismatches: 406
Query Match: 20.04% Indels: 52
DB: 11 Gaps: 18
US-10-090-458-5 (1-1642) x AK085555 (1-3300)
Qy 6 ArgGluValGlyValTTPArgGlnThrArgThrLeuLeuLeuLeuLeuLeuLeuLys 25
Db 167 AGACATTGAGACTGGGTGAGCAATATGGCGCTTCTCTGCAAGAACTGGCTCAGNAAG 226
Qy 26 CysArgThrLysLysSerSerValGlnGluLeuLeuPhePro---LeuPheLeuPhe 44
Db 227 TTCAGGATGAGAGAGAGACCTCTCGAATGGCTTTCTCACTTCTCTTAATATTGTTT 286
Qy 45 TrpLeuLeuLeuLeuSerMetMetHisProAsnLysLysTyrGluGluValProAsnIle 64
Db 287 GTATATCAATTATCTCTCAATTGTCATCAAGTTTCATGCGCTCTGAAATGTCGTGGTG 346
Qy 65 GluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLeuGlyTyrThrProVal 84
Db 1393 TATTTGGTGGCGTTATATTTTACAAAATTTACACTCAGCTAAGTACCGACATCAAGG 1452

```

```

Db 347 GATCTGGACCGTGTGTAACTTTAAAGATTCTTAATATATATGATGGGTGTTGCTCCTGAG 406
Qy 85 ThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeuProAspValIle 104
Db 407 TCCGAGGACCAACCATGAGATATGAACAAAGTGCCTCAGCTCCATTCATGAAGGAAGA 466
Qy 105 IleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLysPheProSer 124
Db 467 ACAATTGGCGCATGTCCTGATCAAAATCCCATGAATGAGCTAGACTTAATTAATCTCGATA 526
Qy 125 AsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuLeuPhe---PhePro 143
Db 527 GATGAGTGAGTGATCATCTTCAAGGACACTTCTCATATCATTTGAATTTTCTGGGGT 586
Qy 144 AspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCys----- 159
Db 587 CAGAGGATCCCTTACACAAAGAGACCAAGACCATTCAGCTTCTTGTGAACCATGAAAC 646
Qy 160 ---SerLysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAla 178
Db 647 AATAAATGATTGGTGAATACTCAGCCCTCTGGGAAAGGGCTTGGTAGCCTTTCAAGCT 706
Qy 179 SerIleAspAlaAlaIleIleGlnLysThrAsnValSerLeuLeuTyrLysGluLeuGlu 198
Db 707 GCCATTAATGCTGGTATCATTCAGATGATCACAACCATCGCGCATGGAAGAGCTGATG 766
Qy 199 SerThr-----LysAlaValIleMetGlyGluThrAlaValAlaGluIle 213
Db 767 TCTGTCAATGGTCAATATGAAGATGCCGCTTTATG-GCCCAAGGAGGAGTGCACACT 825
Qy 214 AspThrPheProArgGlyValIleLeuIleTyrLeuValIleAlaPheSerProPheGly 233
Db 826 GAATTTT-----ATTTCCTCTGTGTTATTCTCTCTCTTCACTTATA 870
Qy 234 TyrPheLeuAlaIleHisIleValAlaGluLysGluLysLysIleLysGluPheLeuLys 253
Db 871 TACTATTATTCAGTCAATATTACTCAGAAAGACAG---TACATCACAACTAATATGCGG 927
Qy 254 IleMetGlyLeuHisAspThrAlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeu 273
Db 928 ATGATGGGGCTTCGAGATCGGCTTCGGGTTTCTCGAGTCGAGATGATGCTGGCTTT 987
Qy 274 IlePheLeuMetSerLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGln 293
Db 988 ATCTTGTGCGTGGCTTGGATGCTCTTATGGTGAATCTGCCAAGTGGTGGTCTTA 1047
Qy 294 SerSerSerIleValIlePheLeuLeuPheLeuTyrGlyLeuSerSerValPhePhe 313
Db 1048 ACTGTTTCATGCTGGTGTCTCTCTCTTTTATGGCGTCTTTGTGATTAACCTTA 1107
Qy 314 AlaLeuMetLeuThrProLeuPheLysSerLysHisValGlyIleValGluPhePhe 333
Db 1108 TCTTTCTCATGAGTGTGTGATAAAGAACCGTTTCTTACTGGCTTGGCTATCTTCATC 1167
Qy 334 ValThrValAlaPheGlyPheIleGlyLeuMetIleIleLeuIleGluSerPheProLys 353
Db 1169 CTTACCGCTTCTGGGGAAGTCTGGGA---TTCACAGCGTTATACAACTCTCTCCTGCC 1224
Qy 354 SerLeuValTrp-----LeuPheSerProPheCysHisCysThrPheValIleGly 370
Db 1225 TTTTGGGAATGGACCTTGTGTTTCTTAGCCCTTT-----GCCTTTACCACTGGA 1275
Qy 371 IleAlaGlnValMetHisLeuGlu---AspPheAsnGluGlyAlaSerPheSerAsnLeu 389
Db 1276 ATGGCCCAAGCTTATACATTTGGATGATGATGATGATGATGATGATGATGATGATGAT 1332
Qy 390 ThrAlaGlyProTyrProLeuIleIleThrIleMetLeuThrLeuAsnSerIlePhe 409
Db 1333 CCAACAACCTCGTACTTAATCATGCCACGCTCTTCATGCTGGTTTGTGATGCTCTCTCTC 1392
Qy 410 TyrValLeuLeuAlaValTyrLeuAspGlnValIle-ProGlyGluPheGlyLeuArgAr 429
Db 1393 TATTTGGTGGCGTTATATTTTACAAAATTTACACTCAGCTAAGTACCGACATCAAGG 1452

```





```
Db 1325 GAGTTTATCTGTGATGGAGAGCCTGTTGATTTCAGATCCCATCTAGTATTGCTATGCA 1384
Qy 1114 ProSerValIleLeuPheThrTyIleAlaSerPheThrPheLysLysIleLeuAsnThr 1133
Db 1385 TCATCTCTCATATTATCATGCTATGTCATTTATTTTCGTAATGCGACAGAAAAAT 1444
Qy 1134 LysGluPheTrpSerPheIleTySerValAlaAlaLeuAlaCysIleAlaIleThrGlu 1153
Db 1445 AGTGGCATCTGGTCAATTTCTTCTTAATGTCACCATCTCTTTATCATCTGCTACTGAC 1504
Qy 1154 IleThr-----PhePheMetGlyTyThrIleAlaThrIleLeuHisTyAlaPheCys 1171
Db 1505 ATAATGAATACAGCGCTTCTAGAGCTGCTAATCTGCACCTTTTAA-----1549
Qy 1172 IleIleIleProIleTyProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTTP 1191
Db 1550 -----GTCCTCTCTTTTCACACTGATTGGCTCACTATTGATTTTTCGAGGTTTCTAT 1603
Qy 1192 LysAsnValArgLysAsnValAspThrTyAsnProTrpAspArgLeuSer-----1208
Db 1604 GACTCTGTG-----GATTACTTGGGAACGTCAGAA 1633
Qy 1209 -----ValAlaValIleSerProTyLeuGlnCysValLeuTTPilePhe 1223
Db 1634 TCTCAACTCGTATTCTGGCATGCTGATACCTTACCTTCTCTTCTTTCTTTTC 1693
Qy 1224 LeuLeuGlnTyTyGluLysLysTyGlyArgSerIleArgLysAspProPhePhe 1243
Db 1694 ATTCGCGATCTCTGGAAGATACCTCAGGAAGAAGTCACTGAGTGGACCTGTGTTTC 1753
Qy 1244 ArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAsp 1263
Db 1754 AGA---ATTCTCCGAGAGT---TGCTCTGCTGTTCCGAACCCAGAGAGCCGAGAA 1807
Qy 1264 GluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCys 1283
Db 1808 GAGGATGAAGATGTTTCAGATGGAAAGAGTGGAAACACGCGCGCCATGCTACCTACAG 1867
Qy 1284 CysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyAspAspLysLys 1303
Db 1868 ACTGATGAACACCATGCTCATTTGCAAGCTGTCTCGGAAGGAATATATAGCGAAGC 1927
Qy 1304 AspPheLeuLeuSerArgLysValLysValAlaThrLysIleTySerPheCysVal 1323
Db 1928 AAAAGGTGCTTTTCCAAAATGAAGAAAAGATGCCACGAGGAACATCTCTCTGTGTT 1987
Qy 1324 LysLysGlyGluIleLeuGlyLeuGlyProAsnGlyAlaGlyLysSerThrIleIle 1343
Db 1988 AAAAAGTGAAGTCTTAGGCTCTTAGGACACATGGAGTGGCAAAAGTACAACAATT 2047
Qy 1344 AsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTySer 1363
Db 2048 AGCATGATAACAGGAGACACAAATACCAACCGCAGGTGAGTGTGTTTG-----2095
Qy 1364 SerGluThrSerGluAspAspAspSerLeuLysCysMetGlyTyCysProGlnIleAsn 1383
Db 2096 ---AAAGGAGCGGTGGGGGTGTCGCCCTCGGATCTTGGGGTACTGCCCTCAGGAGAAC 2152
Qy 1384 ProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyGlyAlaValLysGly 1403
Db 2153 GTGCTGTGGCCCAACCTGACAGTGAAGAACACACCTGGAGTGTGATGCTGTGTGAGGGG 2212
Qy 1404 MetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLys 1423
Db 2213 CTGAAGAAAAGAGTGTGTGGTCCATCACCGTTGTGTGATGGCCCTTAAACCTGCAG 2272
Qy 1424 GluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPhe 1443
Db 2273 GACCATGTGAGGCTCTGTGAGGACCTTGTACAGGGGAGTGAAGAGAAAGCTGTGTTTC 2332
Qy 1444 AlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMet 1463
Db 2333 GTGCTGAGCATCTCTGGGAACCCCTCTGTGTGTTCTTGTGATGAGCCATCGACTGGGATG 2392
```

```
Qy 1464 AspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLys 1483
Db 2393 GACCCGGAAGGCGAGCAGCAAAATGTGCGAGCGCATCCGAGCCACTTTTCACAAACACAGAG 2452
Qy 1484 ArgAlaAlaIleLeuThrThrHisTyMetGluGluAlaGluAlaValCysAspArgVal 1503
Db 2453 CGGGCGCGCCCTCTGACCCACCATTACATGCGAAGCAGAGCGTGTGTGTGACCGGTGTG 2512
Qy 1504 AlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLys 1523
Db 2513 GCCATCATGGGTCTGGGAGGCTGAGGTGCATTGGTTCTATCCAACTCTGAAGAGCAA 2572
Qy 1524 PheGlyLysGlyTyPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeuGluVal 1543
Db 2573 TTGGCAAGATTACCTGCTGAGATGAAGGTGAAG-----ACCCCTCACAGGTG 2623
Qy 1544 AspArgLeuGlnArgGluIleGlnTyIlePheProAsnAlaSerArgGlnGluSerPhe 1563
Db 2624 GAACCCCTGAACACGGAGATCATGAGACTGTTCGCCAGGCTGCTCGCAGAAAGGTAC 2683
Qy 1564 SerSerIleLeuAlaTyLysIleProLysGluAspValGlnSerLeuSerGlnSerPhe 1583
Db 2684 TCCTCTGTGATGCTATTAAGTTGCTTGGAGACGTACGACCTTTATCCGAGCTTTC 2743
Qy 1584 PheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTySerPheSerGlnAla 1603
Db 2744 TCAAAATGGAGAGATTGAAGAAGAAATTTGACCTGGAAGAAATACAGCTTTCACAGTCT 2803
Qy 1604 ThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAspAsnSerCys 1623
Db 2804 ACTCTGGAGAGGCTTCTCTGAGCTCTCCAAGAGCAGAGCTGGAGCTTGGCGAA 2863
Qy 1624 GlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAsp 1638
Db 2864 GAAGCAATTCCTCAGTGAATGGAAGCTCTCCACAGGAAGAG 2908
```

## RESULT 8

```
BC037309 3357 bp mRNA linear HTC 04-MAR-2003
LOCUS Homo sapiens, clone IMAGE:5262251, mRNA.
DEFINITION BC037309
ACCESSION BC037309.1 GI:23336952
VERSION BC037309.1
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3357)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshivuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
Contact: http://www.systemsbio.org
contact: amadansystemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAP Plate: 73 Row: f Column: 14  
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 6005700







```

10349636
2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
  Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  Normalization and subtraction of cap-trapper-selected cDNAs to
  prepare full-length cDNA libraries for rapid discovery of new genes
  Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159
3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
  Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M.,
  Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
  Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
  Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
  Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J.,
  Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  RIKEN integrated sequence analysis (RISA) system--384-format
  sequencing pipeline with 384 multicapillary sequencer
  Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4 The RIKEN Genome Exploration Research Group Phase II Team and the
  FANTOM Consortium.
  Functional annotation of a full-length mouse cDNA collection
  Nature 409, 685-690 (2001)
5 The FANTOM Consortium and the RIKEN Genome Exploration Research
  Group Phase I & II Team.
  Analysis of the mouse transcriptome based on functional annotation
  of 60,770 full-length cDNAs
  Nature 420, 563-573 (2002)
6 (bases 1 to 1529)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
  Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
  Hayaashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
  Horii,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
  Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,W.,
  Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
  Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
  Okazaki,Y., Saito,R., Saito,H., Sakai,C., Sakai,K., Sakazume,N.,
  Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
  Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
  Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
  Muramatsu,M. and Hayashizaki,Y.
  Direct Submission
  Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
  Physical and Chemical Research (RIKEN), Laboratory for Genome
  Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
  RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
  Kanagawa 230-0045, Japan (E-mail:genome-res@gsr.riken.go.jp,
  URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
  Fax:81-45-503-9216)
  cDNA library was prepared and sequenced in Mouse Genome
  Encyclopedia Project of Genome Exploration Research Group in RIKEN
  Genomic Sciences Center and Genome Science Laboratory in RIKEN.
  Division of Experimental Animal Research in Riken contributed to
  prepare mouse tissues.
  Please visit our web site for further details.
  URL:http://genome.gsc.riken.go.jp/
  URL:http://fantom.gsc.riken.go.jp/
  Location/Qualifiers
    1. 1529
      /organism="Mus musculus"
      /mol_type="cDNA"
      /strain="C57BL/6J"
      /db_xref="FANTOM:DB:9430067009"
      /db_xref="MGI:2399173"
      /db_xref="taxon:10090"
      /clone="9430067009"
      /tissue_type="embryonic body between diaphragm region and
      neck"
  source
  /clone.lib="RIKEN full-length enriched mouse cDNA library"
  /dev_stage="12 days embryo"
  61. >1527
  /note="unnamed protein product; ATP-BINDING CASSETTE
  PROTEIN homolog [Homo sapiens] (SPTR|CAB93535, evidence:
  FASTV, 89.4%ID, 76.5%length, match=3750)
  putative"
  /codon_start=1
  /protein_id="BAC28896.1"
  /db_xref="GI:26330332"
  /translation="MEQSLILSETKASSVSTMSLWKKQVSTIAKPHLSLKREKSV
  RAVLLLIIFAVQIFMFFLHSPKNVAVPIKLPDLYFLKPKDPKHYSKTSLLQNS
  TDVINGLIEFPAQNFQVMSAALNVRSEKDYVFSVFNSTWY
  CLPYMNAENHKIKAYTQLSGLPSAYVGVAVDIPLFVFLVILMLGSLPAFHG
  LYFPAKFLAVFCLIAVPSVILFTYASPTFKILNTKEFNSFIYVTLACAVIT
  BITEFLQAVTAVFHYFTCAIPYIPILGLISFKISGWNKMPKENTYINPWRLLVA
  VITFYLQGLIMFLQHYKEKIRKDPFFRALSKAKNKKFPFPPIINDEDEDEDV
  KAERLKVLELMGCCQCCBEKPAIMVNCNLHKSYYDDKKDFLH"
  Alignment Scores:
  Pred. No.: 4,78e-242 Length: 1529
  Score: 2313.00 Matches: 438
  Percent Similarity: 91.34% Conservative: 26
  Best Local Similarity: 86.22% Mismatches: 44
  Query Match: 27.45% Indels: 0
  DB: 11 Gaps: 0
  US-10-090-458-5 (1-1642) x AK034961 (1-1529)
  QY 799 TyrSerValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGlu 818
  Db 1 TATATGTGATTACACAGCAGCCGCGGAGAGAGAAACAGATTCAAAATCTCTTGATGAA 60
  QY 819 MetGluGlnSerLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSer 838
  Db 61 ATGGACAGAGATTACTATTCTCTGAAACGAAAGCTTCTCAGTGAGCACCATTAGC 120
  QY 839 LeuTrpLysGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGlu 858
  Db 121 CTCTGGAGCAGCAAGTGTCTACGATTGCAAGTTTCATTTCTCTCATTGAAACGAGAA 180
  QY 859 SerLysSerValArgSerValLeuLeuLeuLeuLeuPhePheThrValGlnIlePhe 878
  Db 181 AGCAATCAGTCGCGCTGTCTCTCTTTAAATTTTTCAGTTTTCAGATTTT 240
  QY 879 MetPheLeuValHisHisSerPheLysAsnAlaValValProLysLeuValProAsp 898
  Db 241 ATGTTTTTTCCTCATCTCTTTTAAAAATGCTGTGTTCCCATCAAACTCGTTCCAGAC 300
  QY 899 LeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGln 918
  Db 301 TTGTATTTCCTAAAGCCTGGAGATTAACCTCATATAATACAAAACAGCCTGCTCTCAA 360
  QY 919 AsnSerAlaAspSerAspLysSerAspLysPhePheThrSerGlnAsnIleMet 938
  Db 361 AATCTACTGACTCAGATATCAATGCTTATTGAGTTTTCACACACAGACATAATG 420
  QY 939 ValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsn 959
  Db 421 GTGGCAATGTTTAAATGACAGTACTATGTCTCTCTCTCTCTCTCTCTCTCTCAAT 480
  QY 959 ValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyr 978
  Db 481 GTGTGCGTCTGAAAGGACTATGTTTTTCTGCTGTTTCAACAGTACTATGTTAT 540
  QY 979 SerLeuProLysLeuValAsnIleSerAsnTyrTyrLeuTyrHisLeuAsnValThr 998
  Db 541 TGTGTGCGTCAATGATGAAACATCATTAGTAACCTACTATCTTATCATTTAAATGTGACT 600
  QY 999 GluThrIleGlnIleTyrSerThrProPheGlnGlnIleThrAspIleValPheLys 1018

```

```

Db 2863 NNNGAGCATATGAGTATGAGTACCGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2922
Qy 1030 IleIleValThrAlaMetProPheArgPheAlaMetGluAsnAlaGluAsnHisLysIle 1049
Db 2923 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2976
Qy 1050 LysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGln 1069
Db 2977 AAAGCTCATTCCTCCAGCTATGAGTATGAGGCTCTACCCCTCTGCGATACCTGGTGGCCAA 3036
Qy 1070 AlaValAlaAspIleProLeuPhePheIleIleLeuLeuMetLeuGlySerLeu 1089
Db 3037 GCATCGTGGTGAATGTTCTCTGCTATCTTTGATCTCTCTGCTAATG-----CAATAATG 3090
Qy 1090 AlaPheHisTyrGly-----LeuTyrPheTyrThrValLysPheLeuAlaValVal 1106
Db 3091 GATTATATTTTATGAGGAGGAGGATATATTTATATTTCAAAACCTGTTAAATTTCAAACT 3150
Qy 1107 PheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPheThr 1126
Db 3151 CTGTGTAGTATGGCTATGCTCATCTCTGTTCTTGACATATGATGATTCATTTCATT 3210
Qy 1127 PheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyrSerValAlaAlaLeu 1146
Db 3211 TTTGCAATGGGAGAAAAATAGTGGCATTTGGTCAATTTCTTTCTTTAAATTTGGTTCATC 3270
Qy 1147 AlaCysIleAlaIleThrGluIleThr-----PhePheMetGlyTyrThrIleAlaThr 1164
Db 3271 TTTCTGATGTTGCTACTGATCTCTAAATGATATGATGATTTATAGGCTATTTTGGCACC 3330
Qy 1165 IleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLeuIle 1184
Db 3331 ATGTTA-----ATACCTCCCTTCACATTTGATTTGGTCTCTCTATTA 3369
Qy 1185 SerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnProTrp 1204
Db 3370 ATTTTTCTGAG----- 3381
Qy 1205 AspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeu 1224
Db 3382 -----CCTTACCTTCATTTCTCTCATTTTCTTTCTTTCATT 3414
Qy 1225 LeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAspProPhePheArg 1244
Db 3415 CTGCGATCCCTAGAAATGAACCTGCAGGAGAACTAATAGAAAGGATCTGTGTTCCAGA 3474
Qy 1245 AsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAspGlu 1264
Db 3475 ---ATTCTTCCAAGAGC---AACGCTGTTTTCACAAACCAGAGAGCCTGAAGAGAG 3528
Qy 1265 AspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCys 1284
Db 3529 GAGGAAGATATCNNAATGGAAAGAAATGAGACAGTGAATGCTATGCTGNN----- 3576
Qy 1285 GluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAsp 1304
Db 3577 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3636
Qy 1305 PheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSerPheCysValLys 1324
Db 3637 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3696
Qy 1325 LysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThrIleIleAsn 1344
Db 3697 NNNNGTGAAGTTATAGACTGTAGGACACATGAGCTGTGTAAGTCAACTATTAAAG 3756
Qy 1345 IleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSer 1364
Db 3757 ATGATAACTGGGACACAAAAACCACTCAGGACAGGTGATTTTG----- 3801
Qy 1365 GluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnPro 1384

```

```

Db 3802 AAAGCAGCGCGGAGGGAACCCCTGGGCGNNNNNGGGTACTGCCCTCAGGAGAAATGCG 3861
Qy 1385 LeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMet 1404
Db 3862 CTGTGGCCCAACCTCAGACAGTGGCAGCACCTGAGGTGTACGTGGCTGCGTGAAGGGGCTC 3921
Qy 1405 SerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGlu 1424
Db 3922 AGGAAAGGGGACCGCAATGATGCCATCACACGGTTAGTGGATGCGCTCAAGCTCAGAGAC 3981
Qy 1425 HisLeuGlnLysThrValLysLeuProAlaGlyIleLysArgLysLeuLysCysPheAla 1444
Db 3982 CAGCTGAAGGGCTCTCTGTGAGACCTTGTGAGAGGGAATAAAGCGAAAGCTGTGCTCGTG 4041
Qy 1445 LeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAsp 1464
Db 4042 CTGACATCTCTGGGAACCCGTCAGTGTGTTCTTGATGAGCCGTCACCTGGGATGGAC 4101
Qy 1465 ProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLysArg 1484
Db 4102 CCGGAGGGGCGACGACAAATGTGGCGGTGATTCGGGCCACCTTTNNAAACACGAGAGG 4161
Qy 1485 AlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAla 1504
Db 4162 GCGGCCCTCTCTGACCACCCACTACATGCGAGGCTGAGGCGGTGTGTGACCGAGTGGCC 4221
Qy 1505 IleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPhe 1524
Db 4222 ATCATGGTGTCTAGGAAGGCTGAGATGTATTTGTTCCATCAACACCTGAAAGCAAAATTT 4281
Qy 1525 GlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeu---GluVal 1543
Db 4282 GCGAAAGACTACTCTCTGAGATGAAGTGAAG-----AACCTGGCACAAATG 4329
Qy 1544 AspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSerPhe 1563
Db 4330 GAGCCCTCCATCGGAGATCTCTGAGGCTTTTCCCCAGGCTGCTCGGACGAGAGGTTTC 4389
Qy 1564 SerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhe 1583
Db 4390 TCTCTCCCTGATGCTATAAGTTGCGGTTGAGGATGTGCGACCTTTATCATCAGGCTTTC 4449
Qy 1584 PheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAla 1603
Db 4450 TTCAATATAGATAGTATTAACAGAGTTTTCACCTGGAGGAGTACAGCTCTCACAGN 4509
Qy 1604 ThrLeuGluGlnValPheValGluLeuThrLysGluGlnGlu-----GluGlu 1619
Db 4510 ACCCTGGACAGGTTTCTCTGAGGCTCTCCAGGAGCAGGAGCTGGGTGATCTTGAAGAG 4569
Qy 1620 Asp 1620
Db 4570 GAC 4572

RESULT 6
AK034961 1529 bp mRNA linear HTC 18-SEP-2003
LOCUS Mus musculus 12 days embryo embryonic body between diaphragm region
DEFINITION and neck cDNA, RIKEN full-length enriched library, clone.9430067009
product:ATP-BINDING CASSETTE PROTEIN homolog [Homo sapiens], full
insert sequence.
ACCESSION AK034961
VERSION AK034961.1 GI:26330331
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253

```

QY 299 IlePheLeuLeuPhePheLeuTyrglyLeuSerSerValPhePheAlaLeuMetLeuThr 318  
Db 907 GTCTTCACCTCTTCTCCTCTATGGCTGTCTTTGATACTTTAGCTTTCTCCGTGATGAGT 966  
QY 319 ProLeuPheLysLysSerLysHisValglyIleValGluPheValThrValAlaPhe 338  
Db 967 ---GTNN 1023  
QY 339 GlyPheIleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeu 358  
Db 1024 NNN 1083  
QY 359 PheSerProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGlu 378  
Db 1084 NNN 1143  
QY 379 ---AspPheAsnGluGlyAla-----SerPheSerAsnLeuThrAlaGlyProTyr 394  
Db 1144 TATGATGTGAATCTAATAGCCACTTGGATCTCTTCAAAAATCTA-----TAC 1191  
QY 395 ProLeuIleIleThrIleIleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAla 414  
Db 1192 CTCATATAGCTACTCTTTCACTGTGGTTTTCACACCCCTCTGTATTTGGTATTGACA 1251  
QY 415 ValTyrLeuAspGlnValIleProGlyGluPheGlyLeuArgSerSerLeuTyrPhe 434  
Db 1252 TTATATTTTGACAAAATTTTCCCGCTGAATATGACATCGATGTTCTCCCTTGTGTTTC 1311  
QY 435 LeuLysProSerTyrTrpSerLysLysLysLysLysLysLysLysLysLysLysLysLys 454  
Db 1312 CTGAATCTCTGTTTGGTTTCAACACGGAGGGCTAATCATGTGGTCTTGGATGAA 1371  
QY 455 ValAsnGlyAsnIleSerPheSerGluIleIleGluProValSerSerGluPheValGly 474  
Db 1372 ACAGTNTGATCTACACCTAATGACTGTTTGAACACGAGTCTCCAGAAATCTGTGGG 1431  
QY 475 LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnVal 494  
Db 1432 AAGAAGCCATCAGG-----1446  
QY 495 GluAlaLeuArgAsnLeuSerPheAspIleTyrGlyGlyGlnIleThrAlaLeuLeuGly 514  
Db 1447 -----CCAGGTGTGGTGTGTTGACATATATGAGCCAGATCATCGCTCTCTTGGT 1497  
QY 515 HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysProProSer 534  
Db 1498 CACAGTGGAGCTGGAAAAACTACCTGTNAACANNNNNNNNNNNNNNNNNNNNNNNNNNNN 1557  
QY 535 AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
Db 1558 NNN 1617  
QY 555 ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
Db 1618 AGCAAGTNNNGGATTTGTTCNACATCCAATGTGCAATTGATNNNNNNNNNNNNNNNNNN 1677  
QY 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnIleIleGlnGlu 594  
Db 1678 GAANNNNNGGCTGNNNNNAATATAAGGAT-----TTGCCAATGAA 1725  
QY 595 ValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLys 614  
Db 1726 GTGAGAGAA-----1734  
QY 615 LeuSerGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLys 634  
Db 1735 -----GAG 1737  
QY 635 IleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleValTrp 654  
Db 1738 GTTTTGTATTGATGAACGACTGTGCTGATTCGATCTCTTCAAGGCACCGAATATNN 1797

QY 655 AsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPheMetAsp 674  
Db 1798 NATCTCTGAAGAGGGGAATCAGACAGATAATCTCTTCAGACACAGNTTATNNNT 1857  
QY 675 GluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLysCysVal 694  
Db 1858 GAGGCTGACATCTCGCGNACAGGAAGGTTCATATCCAAATGGGAAGCTGAAGTGCA 1917  
QY 695 GlySerSerMetPheLeuLysSerIlystPrglyIleGlyTyrArgLeuSerMetTyrIle 714  
Db 1918 GGCTCTTCTCTGTTCTTAAGAAGAAATGGGCATAGGCTACCATTAAGTTTGCATCTG 1977  
QY 715 AspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAla 734  
Db 1978 AATGAAAGGTGTGATCCAGACAGATATAACATCACTGGTTAAGCAGACATCTCTGATCC 2037  
QY 735 ThrLeuLeuGlnGlnAsnAspGlnLeuValTyrSerLeuProPhePheLysAspMetAsp 754  
Db 2038 AAATTTGACAGACAAAGTGAAGAAACTTGTATATATTTGGCTTTTGGAAAGGACAAAC 2097  
QY 755 LysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyr 774  
Db 2098 AAATTTCCAGNN 2157  
QY 775 GlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGluIle 794  
Db 2158 NNN 2217  
QY 795 AspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluGluMetAspSerLys 814  
Db 2218 GATGAATCAGATATTGGAATTTGGGACAATTACAACTGATGGGCAAAAGATATAGGA 2277  
QY 815 SerPheAspGluMetGluGlnSerLeuLeuLeuSerGluThrLysAlaSerLeuVal 834  
Db 2278 AGCCTTGTGTAGCTGGAACAAGTTTGTCTCTTCCACGAAACAAGGAAACA---ATC 2334  
QY 835 SerThrMetSerLeuTrpLysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThr 854  
Db 2335 AGTGGCTGGCGCTCTGGAGGAGCAGGTCTGTGCAATAGCAAAAGTTCGCTTCTTAAG 2394  
QY 855 LeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeuLeuLeuPhePheThr 874  
Db 2395 TAAAGAAAGAAAGAAAGTCTGTGACTATATATTGCTTTTGGTATTAGCTTTATC 2454  
QY 875 ValGlnIlePheMetPheLeuValHisSerPheLysAsnAlaValValProIleLys 894  
Db 2455 CTCACACTTTTGGAAACATCTATTCTACAGTCATATCAGAAAAGT---TACCCGTGGGA 2511  
QY 895 LeuValProAspLeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSer 914  
Db 2512 CTGTCTCCAATACATACTTCTCTCACCAGACACAACACACAGGATCCTCTGACCCAT 2571  
QY 915 LeuLeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
Db 2572 TTACTGGTCATCAATPAAGACAGGGTCAACCATTTGATTAACATTTTACATCTCAGGCGGA 2631  
QY 935 GlnAsnIleMetValThrMet-----IleAsnAspSerAspTyrValSer 949  
Db 2632 CAAPACATAGCTATAGAAGTGGATGCTCTTGGAACTAGAAATGGCAGACAGATGCCCATCT 2691  
QY 950 ValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAla 969  
Db 2692 -----TACAATGGTGCTATCTATTGTGTGAGTGTATGATAAGNNNNNNNNNNNNNN 2742  
QY 970 AlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsn 989  
Db 2743 NNN 2802  
QY 990 TyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTrpSerTrpPhePhe 1009  
Db 2803 NNN 2862  
QY 1010 GlnGluIleThrAspIleValPheLysIleGluLeuTyrPheGlnAlaAlaLeuLeuGly 1029

```

Qy 1469 nHisMetTrrArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeuTh 1489
Db 4090 GCAATGTGGAGCTATTCAGGCCACCTTTTCAACACAGAACGGGGCCCTCTCTGAC 4149

Qy 1489 rThriHstyMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSerGl 1509
Db 4150 CACCCATTATATGTCAGAGGCTGAGGCTGTGTGTGACCGGTGGCCATCATGTGTCTGG 4209

Qy 1509 yGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyPh 1529
Db 4210 GAGCTGAGGTGTATTGGATCCATCCACATCTCAAAAGCAAAATTTGGCAAGATTACCT 4269

Qy 1529 eLeuGluLysLeuLysAspTrpIleGluAsnLeuGluValAspArgLeuGlnArgGl 1549
Db 4270 GCTGGAGATGAAGTGAAGACC---CTCGAG-----CAGGTGGAGCCCTCAATACGGA 4320

Qy 1549 uileGlnTyrllePheProAsnAlaSerArgGlnGlnLysPheSerSerIleLeuAlaTy 1569
Db 4321 GATCTCAGGCTTTTCCCGAGGCTTTTAGCCAGGACGCTACTCCTCTCTGATGTCFA 4380

Qy 1569 rLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluLual 1589
Db 4381 TAAGTTCGCTCTTGGAGGCTGACGCTTATCCGAGGCTTTCTTCAAAATGGAGAGATT 4440

Qy 1589 aLysHisAlaPheAlaIleGlnGluTySerPheSerGlnAlaThrLeuGlnValPh 1609
Db 4441 TAAACAGAGCTTTGACCTGGAGGAGTACGCTTTACAGTCCACCTCGGAGGCTCTT 4500

Qy 1609 eValGluLeuThrLysGluGlnGlu-----GluGluAspAsnSerCysGlyTh 1625
Db 4501 CCTGGAGCTCTCTAAGGAGCAGGAGCTGGACGGTTTTTGAAGAGGAA----- 4546

Qy 1625 rLeuAsnSerThrLeuTrpTrpGlu 1633
Db 4547 -CTTGATCCCTCAGTCAAGTGGGAAG 4570

RESULT 5
AY414283
LOCUS
DEFINITION Pan troglodytes ABCA9 gene, VIRUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
ACCESSION
VERSION AY414283.1 GI:39770245
KEYWORDS
SOURCE
ORGANISM Pan troglodytes (chimpanzee)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 4620)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS 2 (bases 1 to 4620)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (18-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. 4620
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>4620

gene

```

```

/ gene="ABCA9"
/ locus_tag="HCMS172"

ORIGIN
Alignment Scores:
Pred. No.: 6,52e-242 Length: 4620
Score: 2319.00 Matches: 566
Percent Similarity: 49.79% Conservatives: 251
Best Local Similarity: 34.49% Mismatches: 678
Query Match: 27.52% Indels: 146
DB: 29 Gaps: 28

US-10-090-458-5 (1-1642) x AY414283 (1-4620)

Qy 6 ArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLysAsnTyLeuLeuLys 25
Db 10 AGATCATCATGCGTGGGTGAGCAAAACGTGGGCTCTTCTCTCGAAGAACTGCTCAAAAAA 69

Qy 26 CysArgThrLysLysSerSerValGlnGluLeuPhePro---LeuPhePheLeuPhe 44
Db 70 TGGAGAAATGAAAGANNNNNNNNNNNAATGGGCTCTTTTCATTTCTCTGGTACTGTTT 129

Qy 45 TrpLeuLysLeuLysSerMetMetHisProAsnLysLysTyTrpGluGluValProAsnIle 64
Db 130 CTGTACCTATTCTTCTCAATTATCATCAAGTTTCATGACACTCCTCAAAATGCTTCTCAATG 189

Qy 65 GluLeuAsnProMetAspLysPheThrLeuSerHisLeuLeuGlyTyThrProVal 84
Db 190 GATCTGGGAGCGTGTAGATAGTTTAAATGATCTAATATGTTATGTTATGCTTTCACCTGAA 249

Qy 85 ThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeuProAspValIle 104
Db 250 TCCAAAACTACCAAGAGATAATGAACAAAGTGCTTCAGCCCATTCCTAAAAAGGAAGA 309

Qy 105 IleThrGluGluTyThrAsnGluLysGluMetLeuThrSerSerLysSerProSer 124
Db 310 ACAATCATGGGTGGCCCTGATGAAAAAGCATGATGAATGGATTTGAACTATTCATA 369

Qy 125 AsnPheValGlyValValPheLysAspSerMetSerTyTrpGluLeuArgPhe---PhePro 143
Db 370 GACGAGTGAGAGTGCATCTTTACTGATACCTTCCTACCATTTTGAAGTTTCTTGGGGA 429

Qy 144 AspMetIleProValSerSerIleTyMetAspSerArgAlaGlyCys----- 159
Db 430 CATGAATCCCATGATGAAGACACAGAGACCATTCAGCTCAGTCAAGCAGTGAAT 489

Qy 160 ---SerLysSerCysGluAlaAlaGlnTyTrpSerSerGlyPheThrValLeuGlnAla 178
Db 490 GAAAAAATGAAGTGTGAAGTTCAGAGTCTGGGAGAAAGGCTTTGTAGCTTTTCAAGCT 549

Qy 179 SerIleAspAlaAlaIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGlu 198
Db 550 GCCATTAATGCTGCTATCATGAANNNNNCAACAAATCATTCAGTGTGAGAACAGCTGATG 609

Qy 199 SerThrLysAlaValIleMetGlyGluThrAlaValGluIleAspThrPheProArg 218
Db 610 TCAGTTACTGGTGTACATATGAAGATATACCTTTTGTGCCAAGGAGGAGTGCACACT 669

Qy 219 GlyValIleLeuLysTyLeuValIleAlaPheSerProPheGlyTyTrpLeuAlaIle 238
Db 670 GATTTTTTCATTTCTTTTGGCATTATTTCTTTTTCATACATTATATATATGATCATGTC 729

Qy 239 HisIleValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHis 258
Db 730 AATGTTTACAAAGAAGACAA---TACATTAGCTCATTTGATGATGCTTTCATCTTATCATGCC 786

Qy 259 AspThrAlaPheTrpLeuSerTrpValLeuLeuTyThrSerLeuIlePheLeuMetSer 278
Db 787 GAGTCAGCATTCCTGGCTTTCTCGGGTTTGTATGATGCTTTCATCTTATCATGATGCC 846

Qy 279 LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleVal 298
Db 847 ACTTTAATGGCTCTTATTGTAAATATCGCAAAATTTGCTGCTGCTGCTGCTGTTTGTGAIG 906

```

Db 2158 GCGTTTCCATGACACCCCTGAATGAGGTGTCTCTGAAGTTAGAAG----- 2203  
Qy 794 eAspGluAlaAspTyrSerValPheThrGlnGlnProLeuGluGluMetAspSerLy 814  
Db 2204 -----GATGTTGGAATTTGTGGCGCTTACAAAGCGATGGGCGAAGACATGGA 2253  
Qy 814 sSerPheAspGluMetGluGlnSerLeuLeuLeuLeuLeuSerGluThrLyAsaSerLeuVa 834  
Db 2254 AAGCCTCGTTGAGCTGGAGCAAGTGTGTCTTTA-----GATTCGTGAGAGGCTCAGT 2307  
Qy 834 lSerThrMetSerLeuTrpLysGlnGlnMetTyrThrIleAlaLysPheHisPhePheTh 854  
Db 2308 GAGTGGCATGGCGTGGAGGAGCAGCTCTCGCAGCTGCGAAGTTCGCTTCTCAA 2367  
Qy 854 rLeuLyArgGluSerLysSerValArgSerValLeuLeuLeuLeuLeuLeuPhePheTh 874  
Db 2368 GCTCAAGAATGAACGGAAGAGTCTGATGACTGTGTATTGCTTTTGGCATTAGCTTTGT 2427  
Qy 874 rValGlnIlePheMetPheLeuVal-----HisHisSerPheLysAsnAlaVa 890  
Db 2428 CCCCCAACTTTTGGAAACATCTCTCTACAAGGTATACCAAAAGTTATTTCATGGGGG-- 2485  
Qy 890 lValProIleLysLeuValProAspLeuTyrPheLeuLysProGlyAspLysProHisLy 910  
Db 2486 -----CTATCTCCAAGTATGTATTTCTTTTCCACAGGACAGCCACACAGGA 2532  
Qy 910 sTyrLysThrSerLeuLeuLeuGlnAenSerAlaAspSerAspIleSerAspLeuLeSe 930  
Db 2533 TCCTCTGACCACCTTACTGTGTCATCAATAGACAGGACAAATAATCGATCTTTATACA 2592  
Qy 930 rPhePheThrSerGlnAsnIleMetValThrMet-----I-I 942  
Db 2593 TTCACTGGAAACAACAGACATAGCTTTGGAGTTGATGCTTTGGGACTAGAAATGGCAC 2652  
Qy 942 eAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMethHisSe 962  
Db 2653 CGAGGACAGTCAGTAT-----AATGGAGCTATCATCTCTATCAGGTGA 2694  
Qy 962 rGluLyLe----- 964  
Db 2695 CGAGAGNN 2754  
Qy 965 -----AspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLe 980  
Db 2755 NNN 2814  
Qy 980 uProIleLeuValAsnIleLeSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluTh 1000  
Db 2815 NNN 2874  
Qy 1000 rIleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGl 1020  
Db 2875 CGCATTTTCTGGATACCA----- 2893  
Qy 1020 uLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAl 1040  
Db 2894 -----GTGGCTGGCAGCCTCACCCCTTACATTCG 2922  
Qy 1040 aMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLe 1060  
Db 2923 GATGGGCAGCATCAGTCAGCACCAAAAAAAGTTCTTCCAGCTGTGGACTTCAGGCT 2982  
Qy 1060 uLeuProSerAlaTyrTrpIleGlyGlnAlaValAspIleProLeuPhePheIleI 1080  
Db 2983 TTACCCGCTCTGCTTACTGTGTGGACAGGCTCTGTGGACATTCCTCATATCTTCCTCAT 3042  
Qy 1080 eLeuIleLeuMet-----LeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTy 1098  
Db 3043 TCTCTTCTCATGAATAATAGACAGTGT-----TTTAG 3078  
Qy 1098 rThrValLysPheLeuAlaVal-----ValPheCysLeuIleGl 1111  
Db 3079 TTCTGAGGAGTTTATATCTGTGATGGAGAGCCTGTTGATTTCAGATCCCATGTTAGTATGG 3138

Qy 1111 yTyrValProSerValIleLeuPheThrTyrIleAlaAspPheThrPheLysLysIleLe 1131  
Db 3139 CTATGCATCATCTCTCATATTTCATGATTGATTTCATTTCATTTCGTAATGGCAG 3198  
Qy 1131 uAsnThrLysGluPheTrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIl 1151  
Db 3199 AAAAATAAGTGGCATCTGTCATTTTCTCTTAATTGTCCACCATCTCTTTTATCAVTCG 3258  
Qy 1151 eThrGluIleThr-----PhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAl 1169  
Db 3259 TACTCATATAAATGAATACGGCTTTCTAGACTGCTAATCTGCACCTTTTFA----- 3310  
Qy 1169 aPheCysIleIleIleProIleTyrProLeuLeuGlyCysLeuLeuSerPheIleLysIl 1189  
Db 3311 -----GTCCCTCTCTTTCACACTGATGGCTCACTATGATTTTCTGTAG-- 3355  
Qy 1189 sSerTrpLysAsnValArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSerVa 1209  
Db 3355 ----- 3355  
Qy 1209 lAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTrGl 1229  
Db 3356 -----CCTTACCTTTCATTTCTTTTCTTTTCTTCTGCGATGCTCGGA 3402  
Qy 1229 uLysLysTyrGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLy 1249  
Db 3403 AAGTACCTCAGGAAGATCAGTACGAGTGGACCCCTGTGTTCAGA---ATTTCTCAA 3459  
Qy 1249 sSerLysAsnArgLysLeuProGluProProAspAsnGluAspGluAspVally 1269  
Db 3460 AAAC---AATGATGTGATCAAAACCCAGAGCCAGAGATGAGGATGAAGATGTCA 3516  
Qy 1269 sAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluLysProSe 1289  
Db 3517 AATGAGAGAATAGAAACAGACGAAATGCTTGTGTATCCACAGTTTTCATGAGAACCAGT 3576  
Qy 1289 rIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSerAr 1309  
Db 3577 CATATTCCAAGCTGTCTGCGAAGGAATATATAGGCAGACGAAAGGTGCTTTTCAA 3636  
Qy 1309 gLysValLysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIleLe 1329  
Db 3637 AATGAAGAAAAAGATCGCCACGAGAAACATCTCTCTGT-----GAAGTCCT 3684  
Qy 1329 uGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThrIleLeuAsnIleLeuValGlyAs 1349  
Db 3685 AGGCTCTTAGGACACATGGAGCTGCGAAAGTACAACTAGCATGATATAACAGGAGA 3744  
Qy 1349 pIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAs 1369  
Db 3745 CACAATACCAACCCAGGTCAGGTGTTTTTG-----AAGGAGCGGTGG 3789  
Qy 1369 pAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTrpProAspTh 1389  
Db 3790 GGGTGTGCTCCCTCGGATTCCTGGGTATCTGCCCTCAGGAGAACGTGCTGTGGCCCACT 3849  
Qy 1389 rThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAspMe 1409  
Db 3850 CACAGTGAAGAAACACCTGGAGTTGATGCTGTCTGAAGGGCTCGAAGAAAAGGATGC 3909  
Qy 1409 tLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysTh 1429  
Db 3910 TGTGGTCACCATCACCGTTTGGCAGATGGGCTCAAGCTGCGAGCACCGAGCTGAAGTCTCC 3969  
Qy 1429 rValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGl 1449  
Db 3970 TGTGAAGACGCTGTACAGAGGAGTGAAGAGAAAGCTGTGTGTGTGTGATCCTCTGGG 4029  
Qy 1449 yAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGl 1469  
Db 4030 GAACCCCTCTGTGTTACTTCTGTGATGAGCCATCCACCGGATGAGACCCCGAGGGCAGCA 4089





```
Db 3409 TTCATTGCGTCCCTAGAAATGAACCTCAGGAGAACTATAGAGAAAGGATCTGTG 3468
Qy 1243 PheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGlu 1262
Db 3469 TTCAGA---ATTTCCTCAAGAAGC---AACGCTATTTTCCAAACCCAGAGAGCTGAA 3522
Qy 1263 AspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGln 1282
Db 3523 GGAGAGGAGAGAGATATCCAGATGGAAGAATGAGACAGTGAATGCTATGGCTGTGCGA 3582
Qy 1283 CysCysGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLys 1302
Db 3583 GACTTTGATGACACACCCGTCATCATTCGCGAGCTGTCTACGGAAGGAATATGACAGCAAA 3642
Qy 1303 LysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCys 1322
Db 3643 AAGAAAATTCCTTTCTAAAGGAGAGAAAATTCACCAAGAAATGTCCTTTTGT 3702
Qy 1323 ValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThrIle 1342
Db 3703 -----GAAGTTATAGGACTCTTAGGCACAACTGAGCTGGTAAAGTCAACT 3750
Qy 1343 IleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyr 1362
Db 3751 ATTAAGATGATACTGGAGACACAAACCACTGCAGGACAGGTGATTTTG----- 3801
Qy 1363 SerSerGluThrSerGluAspAspSerLysCysMetGlyTyrCysProGlnIle 1382
Db 3802 -----AAGGGAGCGGTGGAGGGAACCCCTGGCTTCCTGGGGTACTGCCCTCAGGAG 3855
Qy 1383 AsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLys 1402
Db 3856 AATGCGTGTGGCCCACTGACAGTGAAGCAGCAGCAGTGGAGTGAAGTGGCGGTGAA 3915
Qy 1403 GlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeu 1422
Db 3916 GGCTCAGGAAGAGGGAGCGCAATGATCGCCATCACAGGTAGTGGATGGCTCAAGCTG 3975
Qy 1423 LysGluHisIleuGlnLysThrValLysLeuLeuProAlaGlyIleLysArgLysLeuCys 1442
Db 3976 CAGGACAGGTGAAGGCTCCCGTGAAGACCTTGTGAGAGGGAATAAGCGAAAGCTGTGC 4035
Qy 1443 PheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGly 1462
Db 4036 TTGTGCTGAGCATCTCTGGGAACCCGTAGTGGTGTCTTGTGATGAGCGTGCACCGGG 4095
Qy 1463 MetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArg 1482
Db 4096 ATGACCCCGAGGGCGCAGCAAAATGTGGCAGGTGATTCGGGCCACCTTTAGAAACAG 4155
Qy 1483 LysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCysAspArg 1502
Db 4156 GAGAGGGCGCCCTCTCTGACACCCACTATGCGAGAGCTGAGCGGTGTGTGACCGA 4215
Qy 1503 ValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLysSer 1522
Db 4216 GTGGCCATCATGTGTGAGGAGGCTGAGATGATTGTTCCATCCAAACACCTGAAAGC 4275
Qy 1523 LysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeu--- 1541
Db 4276 AATTGCGAAAGACTACCTCGTGGAGATGAAGCTGAAG-----AACCTGGCA 4323
Qy 1542 GluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGlu 1561
Db 4324 CAAATGGAGCCCTCCATGTCAGAGATCTCTGAGGCTTTTCCCCAGGCTGCTCAGCAGGA 4383
Qy 1562 SerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGln 1581
Db 4384 AGGTCTCTCCCTGATGGTCTATAAGTTCCTGTTGAGGATGTGCGCTTTATCACAG 4443
Qy 1582 SerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSer 1601
Db 1601
```

```
4444 GCTTTCTTCAAATTAGAGATAGTTAAACAGAGAGTTTCGACCTGGAGAGTACAGGCTCTCA 4503
Qy 1602 GlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGlu----- 1617
Db 4504 CAGTCTACCTCCCTGGAGCAGGTTTCTCGAGCTCTCCAGGAGCAGGAGCTGGGTGATCTT 4563
Qy 1618 GluGluAsp 1620
Db 4564 GAAGAGGAC 4572

RESULT 4
AY414284
LOCUS AY414284
DEFINITION Mus musculus ABCA9 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY414284
VERSION AY414284.1 GI:39770246
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 4594)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 4594)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submision
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..4594
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>4594
/gene="ABCA9"
/locus_tag="HCM5172"
gene
ORIGIN
Alignment Scores:
Pred. No.: 4,36e-293 Length: 4594
Score: 2785.00 Matches: 648
Percent Similarity: 54.65% Conservative: 275
Best Local Similarity: 38.37% Mismatches: 537
Query Match: 33.05% Indels: 229
DB: 29 Gaps: 35

US-10-090-458-5 (1-1642) x AY414284 (1-4594)
Qy 6 ArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLysAsnTyrLeuIleLys 25
Db 10 AGACACTTGACATGGGTGCGCAAAATGGGCGCTTCTCTGCAAGAACTGGCTCAGAAAG 69
Qy 26 CysArgThrLysSerSerValGlnGluIleLeuPhePro---LeuPhePheLeuPhe 44
Db 70 TTCAGATGAGAGAGAGAGACCTCTCTGAGTGGGCTTTTCTCACTTCTTAAATATGTTT 129
Qy 45 TrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGluValProAsnIle 64
Db 130 GTATATCAATTATCTCCAAATTTCATCAAGTTTCATGACGCTCTCGAAATGTCGTGGTG 189
```

Qy	515	HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysProProSer	534
Db	1498	CACAGTGGAGCTGGAAAAACATACCCTGTAAACATACTTAGTGGGTGTGAGTTCCCAACA	1557
Qy	535	AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla	554
Db	1558	TCAGGTTCACTCACTGCTCTATATCAACACTTTCAGAGTGGCTGATATAGAAAATATC	1617
Qy	555	ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu	574
Db	1618	AGCAAGTTCACCTGGAATTTGTGCCACAATGCAATTTGGAATTTCTCACTGTGAAA	1677
Qy	575	GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGlu	594
Db	1678	GAAACCTCAGGCTGTTTGTAAATAAAGGGATT-----TTGCCACATGAA	1725
Qy	595	ValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLys	614
Db	1726	GTGAGAAA-----	1734
Qy	615	LeuSerGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLys	634
Db	1735	-----GAG	1737
Qy	635	IleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleValTrp	654
Db	1738	GTTTGTGATTTGGATGAACACGACTGCTGGATTTGGATCCTCTTCAAGGCACCGAATATGG	1797
Qy	655	AsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPheMetAsp	674
Db	1798	AATCTCTGAAAGAGGGGAAATCAGACAGAGTAATCTCTTCAGCACCCAGTTTATAGAT	1857
Qy	675	GluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLysCysVal	694
Db	1858	GAGCTCACATCTGGCGGACAGAAAGGTTCATCCAAATGGGAAGCTGAAGTGGCA	1917
Qy	695	GlySerSerMetPheLeuLysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIle	714
Db	1918	GGCTCTCTCTGTTCTTAAGAGAAATGGGCATAGGCTACCATTTTAAGTTTGCATCTG	1977
Qy	715	AspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAla	734
Db	1978	AATGAAGGTGTGATCCAGAGAGTATAAACATCACTGTTAAGCAGCACATCTCTGATGCC	2037
Qy	735	ThrLeuLeuGlnAsnAspGlnGlnLeuValTyrSerLeuProLysAspMetAsp	754
Db	2038	AAATTGCAGCACAAAGTGAAGAAAACTTGATATATTTTGCCTTTGGAAAGGACAAAC	2097
Qy	755	LysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyr	774
Db	2098	AAATTTCCAGNNNNNNNNNNNNNNNNNNNNNTGTTCTAACAGGCATTGAGGATTAT	2157
Qy	775	GlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGluIle	794
Db	2158	GGTGTTTCCATAACAATTTGAATGAGGTGTTTCTGAAATTAGAAGGAAATCAACTATT	2217
Qy	795	AspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluMetAspSerLys	814
Db	2218	GATGAATCAGATATTGGAATTTGGGACATTAACAACGATGGGGCAAAAGATATAGGA	2277
Qy	815	SerPheAspGluMetGluGlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuVal	834
Db	2278	AGCCTTGTGAGCTGGAAACAAGTTTCTCTCTCTCCACGAAACAGAGAAAAACA--ATC	2334
Qy	835	SerThrMetSerLeuTyrLysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThr	854
Db	2335	AGTGGCGTGGCTCTGGAGCAGCAGGTCTGTGCAATAGCAAAAGTTCGCTTCCCTAAAG	2394
Qy	855	LeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeuLeuIlePhePheThr	874
Db	2395	TTAAAGAAAGAAAGAAAAAGCCTGTGGACTATATTATTTGCTTTTGTGTAATAGCTTTATC	2454

875 ValGlnIlePheMetPheLeuValHisHisSerPheLysAsnAlaValValProIleLys 894  
 2455 CCTCAACTTTTGGACACATCTATTCTCAGAGTCAATATCAGAAAAGT---TACCCTGGGAA 2511  
 895 LeuValProAspLeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSer 914  
 2512 CTGCTCCAAATACATACATCTCTCTCACAGGACAAACACACAGGATCTCTGACCCAT 2571  
 915 LeuLeuLeuGlnHisSerAlaAspSerAspIleSerAspLeuLysSerPhePheThrSer 934  
 2572 TTACTGTGTCATCAATAGACAGGGTCAACCATGTAACTTTTATCATTCACCTGAGCGGA 2631  
 935 GlnAsnIleMetValThrMet-----IleAsnAspSerAspTyrValSer 949  
 2632 CAGAAACATAGCTATAGAAGTGGATGCTTTGGAACTAGAAATGGCACAGATGACCCATCT 2691  
 950 ValAlaProHisSerAlaAlaLeuAsnValMetHisSerGlnLysAspTyrValPheAla 969  
 2692 -----TACATGGTGTCTATCTGTCAGGTGATGAAAGNNNNNNNNNNNNNNNNN 2742  
 970 AlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsn 989  
 2743 NNN 2802  
 990 TyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTyrPheThrProPhePhe 1009  
 2803 NNN 2862  
 1010 GlnGluIleThrAspIleVal-----PheLysIleGluLeuTyrPheGlnAlaLeu 1027  
 2863 NNGAGCATATGATTATGATGATGGGTACCGAAGTAACACCTTCTCTGGATACCGATG 2922  
 1028 LeuGlyIleIleValThrAlaMetProProTyrPheAlaValGluAsnAlaGluAsnHis 1047  
 2923 -----CGAGCTCTTTCACTCCATCATGTCATGTCAGCAGCATGGTGACTAC 2970  
 1048 LysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIle 1067  
 2971 AAAAAAAGCTCATTCACAGCTACGAGTTTCAGGCTCTACCTCTCTGCATCTGGTTT 3030  
 1068 GlyGlnAlaValValAspIleProLeuPhePheIleIleLeuLeuMetLeuGlySer 1087  
 3031 GGCCAGCATGTTGGATGTTCCCTGTACTTTTGATCTCTCTGCTAATG-----CAA 3084  
 1088 LeuLeuAlaPheHisTyrGly-----LeuTyrPheTyrThrValLysPheLeuAla 1104  
 3085 ATAATGATTATATATTTTAGCCAGAGAGATATATTTATATTCAAAACCTGTTAAT 3144  
 1105 ValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSer 1124  
 3145 CAAATCCGTGTAGTATTGGCTATGTCATCTCTGTTTCTTGACATATGATGATTCA 3204  
 1125 PheThrPheLysIleLeuAsnThrLysGluPheTrpSerPheIleTyrSerValAla 1144  
 3205 TTCATTTTTCGCAATGGGAGAAAATAAGTGGCATTTGTCATTTTCTTCTTAATATGG 3264  
 1145 AlaLeuAlaCysIleAlaIleThrGluIleThr-----PhePheMetGlyTyrThrIle 1162  
 3265 GTCACTCTTCGATAGTTGCTACTGATCTAAATGAATATGGATTTCTAGGCTATTTTT 3324  
 1163 AlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCys 1182  
 3325 GGACACCATGTA-----ATACCTCCCTTCATATGATGGCTCT 3363  
 1183 LeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsn 1202  
 3364 CTATTCATTTTCTCAG----- 3381  
 1203 ProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIle 1222  
 3382 -----CCTTACCTCATTTCTCATATTTTCTT 3408  
 1223 PheLeuLeuGlnTyrTyrGlnLysLysTyrGlyIleArgSerIleArgLysAspProPhe 1242



Db 1747 GTCAATTGCATATATGAAGGTGAGATTACTGTCACCTGCTGGGCCACAGTCGGAACAGAAA 1806  
Qy 520 sSerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
Db 1807 AAGCACACTGATGAATATCTGTGTGAGCTGTGTCCACCCCTCTGATGGGTTCCTCTAT 1866  
Qy 540 eTyGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db 1867 ATATGGACACAGAGCTCTCTGAATAGATGAATGTTTGAAGCAAGAAAATGATTGGCAT 1926  
Qy 560 eCysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLe 580  
Db 1927 ATGTCCCGCAGTCAGATATAAACTTTGATGTTCTGCACAGTAGAAGAAAATTTATCAATTTT 1986  
Qy 580 uAlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLe 600  
Db 1987 GGCCTTCAATCAAGGAATACCGCCACACATATAATTCAGAAAGTCGAGAAAGTGTGCT 2046  
Qy 600 uAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLy 620  
Db 2047 GGATCTGGACATGCAAGCCATCAAGATAATCAAGCGAAAAGTTAAGCGGTGGTCAGAA 2106  
Qy 620 sATcLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGl 640  
Db 2107 AAGGAAGCTGTCTGTAGAAATTCAGATTCTCGGGAATCCAAAGATACCTCTGCTAGACGA 2166  
Qy 640 uProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyAr 660  
Db 2167 GCCTACAGCAGGAATGACCCCTCTCTCGCATATTTGTTTGGAACTCTTCAAGATATAG 2226  
Qy 660 GlyAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAl 680  
Db 2227 AAAGGCTAAACAGAGTACCGGTGTTAGTACTCTCACTTCATGGATGAGGCTGACATTTCTG 2286  
Qy 680 aAspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLe 700  
Db 2287 CGACAGGAAGCTGTCTATACCAAGGAATGCTGAGTGTGTTGTTCTTCAATTTTCT 2346  
Qy 700 uLysSerLysTrpGlyIleGlyTyArgLeuSerMetTyIleAspLysTyCysAlaIth 720  
Db 2347 AAAGAGTAATGGGGAATCGGCTACCCCTGAGCATGTATATAGACAGATACCTGTGCCAC 2406  
Qy 720 rGluSerLeuSerSerLeuValLysGlnHisIleProGlyValAlaThrLeuLeuGlnAs 740  
Db 2407 AGATCACTGTCTGCTGTTAGCGACGACATACCCGCGCGCTACTGTCAGCAGAA 2466  
Qy 740 nAspGlnLeuValTySerLeuProPheLysAspMetAspLysPheSerGlyLeuPh 760  
Db 2467 TGACCAGCAGCTCTGTACAGCCTGCTTTCAAGACATGACAAAATTTTCAGGCTGT 2526  
Qy 760 eSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyGlyValSerMetThrTh 780  
Db 2527 TTTCTGCTTAGACATTCAATCAAACTGGGTGTTATTTCTTATGGTGTTCATGACAC 2586  
Qy 780 rLeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrs 800  
Db 2587 ATTGGAAGATGATTCTTAAAGCTAGAAGTTGAAGCAGAGAAAATTTGACCAAGCAGACTATAG 2646  
Qy 800 rValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGl 820  
Db 2647 TGTATTTACAGCGCGCGGAGGAGAAACAGATTCAAAATCCCTTTGTGTAATGGA 2706  
Qy 820 uGlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTr 840  
Db 2707 ACAGATTTACTATTCTCTGAAACGAAAGCTTCTTCAGTGAGCACCATTGAGCCTCTG 2766  
Qy 840 PlysGlnLeuMetTyThrIleAlaLysPheHisPheThrLeuLysArgGluSerLy 860  
Db 2767 GAAGCAGCAAGTGTCTAGATGCAAAAGTTTCATTTCTCTCATTTGAACAGAGAACAA 2826  
Qy 860 sSerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPh 880  
Db 2827 ATCAGTGGCGCTGTGTGCTTCTGCTTTTAAATTTTTTTTGGAGTTTTCAGATTTTATGTT 2886

Qy 880 eLeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTy 900  
Db 2887 TTTTCTCATCATTTCTTTTAAATAGTGTGGTTCCTCATCAACTCGTTCACAGACTTGTA 2946  
Qy 900 rPheLeuLysProGlyAspLysProHisLysTyTyLysThrSerLeuLeuLeuGlnAsnSe 920  
Db 2947 TTTCTAAAGCTGGAGATAAACCTCATAAATACAAACCAAGCCTGCTGCTTCAAAATTC 3006  
Qy 920 rAlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValTh 940  
Db 3007 TACTGACTCAGATATCAATGTCCTTATTTAGTCTTTTGGCACACCAAGAACATATGGTGGC 3066  
Qy 940 rMetIleAsnAspSerAspTyValSerValAlaProHisSerAlaAlaLeuAsnValMe 960  
Db 3067 AATGTTTAAATCAGAGTACTATGTGTCTGCTGCTCCTCACAGTGGGTCTCTCAATGGT 3126  
Qy 960 tHisSerGluLysAspTyValPheAlaAlaValPheAsnSerThrMetValTySerLe 980  
Db 3127 CGGTTCTGAAAGGACATATGTTTTTCTGCTGTTTTTCAACAGTACTATGTTTTATGTTT 3186  
Qy 980 uProIleLeuValAsnIleIleSerAsnTyTyLeuTyHisLeuAsnValThrGluTh 1000  
Db 3187 GCCAGTCATGATGAACATCAATAGTAATCTATCTTTATCATTTAAATGTGACTGAAGC 3246  
Qy 1000 rIleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGl 1020  
Db 3247 CATCCAGACCTGGAGTACCCGTTTCATTCAGAAATTTACTGACATTTGTTTTAAATGA 3306  
Qy 1020 uLeuTyThrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyThrPheAl 1040  
Db 3307 GCTATATTTCAAGCAGCTTTGCTTGGAAATCATTTGTAATGTCGAATGCCACTTACTTTC 3366  
Qy 1040 aMetGluAsnAlaGluAsnHisLysIleLysAlaTyThrGlnLeuLysLeuSerGlyLe 1060  
Db 3367 CATGAAATGCAGAGATCATAGATCAAGCTTACTCAACTTAACTTTCGGGTCT 3426  
Qy 1060 uLeuProSerAlaTyTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleI 1080  
Db 3427 TTTGGCATCAGCTTATGGGTGGACAGCTGTGTGGAATATCCCTGTTTTTTGTTGT 3486  
Qy 1080 eLeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyGlyLeuTyThrTyThrVa 1100  
Db 3487 TCTGATTTGATGCTGGGAAGTTTATTTGCATTTTCATCATGAGTGTATTTTATCCTGC 3546  
Qy 1100 lLysPheLeuAlaValValPheCysLeuIleGlyTyValProSerValIleLeuPheTh 1120  
Db 3547 AAAATTTCTGCTGTGGTGTGTTTGGCTCATTTGCTTATGTGCTTCCTCATCTTCTTAC 3606  
Qy 1120 rTyIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheI 1140  
Db 3607 GTACATAGCTTCGTTCACTTTCAAGAAAATTTTAAATACCAAGGAATTTTGTCTATTAT 3666  
Qy 1140 eTySerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPheMetGlyTy 1160  
Db 3667 CTATTTCTGTGACCGCATTTGGCTTGTGGCAATCAGGNAACAACATTTCTTCTTGCAATA 3726  
Qy 1160 rThrIleAlaThrIleLeuHisTyAlaPheCysIleIleIleProIleTyProLeuLe 1180  
Db 3727 TCCAGTACCGGTGCTTTTTCATTCACCTTCTGCATAGCCATTCCAATCTACCTCTCTCT 3786  
Qy 1180 uClyCysLeuLeuSerPheIleLysIleSerTyIlysAsnVal 1194  
Db 3787 GGGTTGTCTGATTTCTTTTCAAAAAGGGTCTTGGAGAAGTAATG 3829

## RESULT 3

AY414282

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

4620 bp DNA linear GSS 17-DEC-2003  
Homo sapiens ABCA9 gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.  
AY414282  
AY414282.1 GI:39770244  
GSS.

**JOURNAL**

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

**COMMENT**

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.  
URL: <http://genome.gsc.riken.go.jp/>  
URL: <http://fantom.gsc.riken.go.jp/>.

```

FEATURES
  source
    location/Qualifiers
      1. .3831
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6J"
        /db_xref="PANTOM_DB:B230352G14"
        /db_xref="MGI:2411591"
        /db_xref="taxon:10090"
        /clone="B230352G14"
        /sex="male"
        /tissue_type="corpora quadrigemina"
        /clone_lib="RIKEN full-length enriched mouse cDNA library"
        /dev_stage="adult"
      1. .3831
        /note="ATP-BINDING CASSETTE PROTEIN homolog [Homo sapiens]
        (SPT)|CAB93535, evidence: FASTA, 89.4%ID, 76.5%length,
        match=3750)"
    misc_feature

```

```

match=3750)

ORIGIN

Alignment Scores:
Pred. No.:          0
Score:              5571.00
Length:             3831
Percent Similarity: 94.90%
Matches:            1076
Best Local Similarity: 90.04%
Conservative:       58
Mismatchatches:     60
Query Match:        1
Indels:              1
DB:                  0
Gaps:                0

US-10-090-458-5 (1-1642) x AX046203 (1-3831)

```

```

US-10-090-458-5 (1-1642) x AK046203 (1-3831)

Qy      1 MetSerThrAlaIleArgGluValGlyValTrrpGlnThrArgThrLeuLeuLeuLys 20
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     247 ATGGCTACTGCAATTAGGAGATGCGAGTTGGAGACAGACCAGAACTTCTACTGAAA 306

Qy     21 AsnTyrLeuIleLysCysArgThrLysSerValGlnGluIleLeuPheProLeu 40
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     307 AATTACCTAATTAATGCAGACCTAAATAAGTAGTGTTCCAGAAATCTTTTCTCTCA 366

Qy     41 PhePheLeuPheTrrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     367 TTTTCTTCTATTTGGCTGATATTAGTTAGCATGATGCATCCAAATAAGAAATATGAAG 426

Qy     61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     427 GTATCTGTATATAGACTCAGCCCTATGCACAAATTCACGCTTCCAAACGTTATCTTGA 486

Qy     81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     487 TACACTCCCGTGACCTAACATTACACAGCAGCATTTATGCAGAGGGTTTTCACGATCATCT 546

Qy     101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     547 CCCAAGGTTTATAGTTACTTGAAGAATACCCAATCAGAGAAAGAACTGGTAGCGCAAGTCCT 606

Qy     121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     607 TCTAAGTCCAGCAACTTCGTAGGTGGTGTCTTCAAGACACCACTGCTCATGAACCTCGT 666

Qy     141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160

```

[illegible]



Db TCTAAGTCCAGCAACTCGTAGGTGTTGGTTTTCGAAGACACCATGTCTTATGAATTCGT 620  
Qy PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
Db TTTTCTTCTGAATGATTCAGGTGCTCTCTATTTATGATTAATCAAGAGGCTGTTCA 680  
Qy LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
Db AAGACATGTGATGCTGCTCAGTACTGCTTTGGGGTTTACAGTTCTCGCAGGCATCGATA 740  
Qy AsnAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
Db GATGCTGCCATTATACAGCTGAGACCAATGTTCTGTGTGGAGCGAGCTGGAGTCGACC 800  
Qy LysAlaValIleMetGlyGluThrAlaValAlaGluIleAspThrPheProArgGlyVal 220  
Db AAAGCTGTGATCATGGAGAGCGGCTGTTGTGGAGATTGACACCTTCCCGGAGGGTC 860  
Qy IleIleuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheIleuAlaIleHisIle 240  
Db ATCCCTCATCTACCTCGTCATAGCCTTCTCGCCCTTCGGCTACTTCTCGCAATCCACATC 920  
Qy ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
Db GTGCGAGAAAAGAAAGAGTTAAAGGAATTTTAAAGATAATGGGACTTCATGACACT 980  
Qy AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerIleuIlePheLeuMetSerLeuLeu 280  
Db GCTTTTGGCTTTCCCTGGGTCTCTCTGACGCAAGCTTGATTTTCCTTATGTCCTCGCTT 1040  
Qy MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300  
Db ATGGCTGTCTATCGACACAGCTTCTCGTTATTCCTCAGAGTAGACGATTTGTGATCTTT 1100  
Qy LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
Db CTACTGTCTCTTATATGATGATGTCATCTGTGTTTTTGGCTTTAATGTCTGACGCTCTT 1160  
Qy PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
Db TTTTAAATAATCAAAACACGTGGAGTCGTTGAGTTTTTTGTACCCGTGGTGTGGATTT 1220  
Qy IleGlyLeuMetIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
Db GTTGGCCTGCTGATGTCTCATAGAAAGTTTCCCGAGGTGCTGTGTGTGCTCTTCACT 1280  
Qy ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
Db CCTTTGTGTCAGTGTGCTTCTGATTTGGGATTTGCACAGTCAATGATTTAGAGATTTT 1340  
Qy AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
Db AATGAGGTCCTTATTTTCTAATTTGACTGAAGTCCCTATCTCTAATTTACTATT 1400  
Qy IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
Db ATCATGTAGCTTGTGACATGTGTCTATGTCCTCTGCTGTGTATCTCGACCAAGTC 1460  
Qy IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
Db ATTCCAGGGGAATTCGCTTGAGGAGGTCACTTTGTATTTTGAAGCCATCGATTGG 1520  
Qy SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
Db TCAAAAACAAAAGAAACTATAAGAGGCTATCGAGGCAACATTAATGGCAATATTAGT 1580  
Qy PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
Db CTCAATGAATTTGTTGAGCCCGTTTCTTCAGAAATTTATAGGAAAGAGCTATPAGAAAT 1640  
Qy SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAsnValGluArgAsnLeu 500

Db AGTGGTATTCAGAAATCCTATAGAAAGAAACTGAGAACGTGGAGGCTTTGAGAAATTTG 1700  
Qy SerPheAspIleTyrGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
Db TCATTTGACATATATGAAGGTACAGTTACTGCACTGCTGGCCACAGTGGACAGGAAA 1760  
Qy SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
Db AGCACATGATGAATATTTCTGTGTGGACTGTGTCCACCTCTGATGGTTTGTCTTCATA 1820  
Qy TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db TATGGACACACAGTCTCTGAATATAGATGAATGTTTGAAGCAAGAAAATGATTTGGCATA 1880  
Qy CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
Db TGTCCGAGTCAGATATAAACTTTTGTCTGACAGTAGAAGAAAATTTATCAATTTTG 1940  
Qy AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
Db GCTTCATCAAGGAATACACGCCCAACATATATTTCAAGAGTGCAGAAAGTGTTCGTG 2000  
Qy AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
Db GATCTGGACATGCAAGCCATCAAGAGATAATCAAGCGAAAAGTTAAGCGGTGGTCAGAAA 2060  
Qy ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
Db AGGAGCTGTCTGTAGGAATTTGCAAGTTCTCGGAAATCCAAAGATACCTCTCTAGACGAG 2120  
Qy ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
Db CCTCAGCAGGAATGGACCCCTGCTCGCCATATTTGTTGGAATCTTTAAAGTATAGA 2180  
Qy LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db AAGCTTAACAGAGTGCACCGTGTCTGACTCCTCATGTGATGAGGTGACATTTCTTGGCC 2240  
Qy AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db GACAGGAAGCTGTCTATACAGGAATGCTGAAGTGTGTGTTCTTCAATTTTCTTA 2300  
Qy LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db AAGAGTAAATGGGAATCGGCTACCGCTGACGATGTATATAGACAGATCTGTGCCACA 2360  
Qy GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db GAGTCACCTGCTGCTGCTGTTAGCAGCACATATCCCGAGCCGCGCTACTGACAGCAAT 2420  
Qy AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db GACCAGCAGCTCGTGTACAGCTGCCCTTCAAGACATGACAAATTTTCAGGCTTCTTT 2480  
Qy SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db TCTGCTCTAGACATTCATTCAACCTGGGTGTTATTTCTTATGTGTGTTTCATGACACA 2540  
Qy LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db TTGGAAGATGATTTCTTAAAGCTAGAAGTTGAACAGAAATTTGACCAAGCAGACTATAGT 2600  
Qy ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db GTATTACACAGCAGCCGCGGAGGAAGAAACAGATTCAAAATCTTTTGTATGAATGGAA 2660  
Qy GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db CAGAGTTTACTTATCTCTGTAACGAAAGCTTTCTTCAAGTACCATCAGGCTCTGG 2720  
Qy LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
Db AAGCAGCAAGTGTCTACGATTGCAAGTTTTCATTTCTCTCTCATTTGAAACGAGAAAGCAA 2780





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 12, 2004, 21:03:35 ; Search time 5792 Seconds  
(without alignments)  
8465.763 Million cell updates/sec

Title: US-10-090-458-5  
Perfect score: 8426  
Sequence: 1 MSTAIREVGWQRTRILLK.....CGTLNLTWERTQEDRVVF 1642

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_17USP10.spool/US10090458/runat\_08042004.121553\_25490/app.query.fasta\_1.1759  
-DB=EST -QMT=fastcap -SUFFIX=ret -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEA=ext -MINLEN=500 -MAXLEN=2000000000  
-USER=US10090458 @CNG 1.1 3643 @runat\_08042004.121553\_25490 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : EST:  
1: em\_estbun.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hcc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hcc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_vrt.\*  
22: em\_gss\_vrt.\*  
23: em\_gss\_vrt.\*  
24: em\_gss\_vrt.\*  
25: em\_gss\_vrt.\*  
26: em\_gss\_vrt.\*  
27: em\_gss\_vrt.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5815	69.0	3950	11 AK047188	AK047188 Mus muscu
2	5571	66.1	3831	11 AK046203	AK046203 Mus muscu
3	2846	33.8	4620	29 AY414282	AY414282 Homo sapi
4	2745	33.1	4594	29 AY414284	AY414284 Mus muscu
5	2319	27.5	4620	29 AY414283	AY414283 Pan trogl
6	2313	27.5	1529	11 AK034961	AK034961 Mus muscu
7	2024.5	24.0	4119	11 AK029256	AK029256 Mus muscu
8	1837	21.8	3357	11 BC037309	BC037309 Homo sapi
9	1688.5	20.0	3300	11 AK085555	AK085555 Mus muscu
10	1677	19.9	1346	11 U66872	U66872 Homo sapien
11	1471	17.5	6600	29 AY405473	AY405473 Mus muscu
12	1444	17.1	6606	29 AY405471	AY405471 Homo sapi
13	1427	16.9	850	14 CD250953	CD250953 AGENCOURT
14	1396	16.6	6606	29 AY405472	AY405472 Pan trogl
15	1382	16.4	2844	11 AK031843	AK031843 Mus muscu
16	1351.5	16.0	872	12 EG564439	EG564439 602584385
17	1270	15.1	771	12 EG564439	EG564439 602584385
18	1251	14.8	809	11 BC029426	BC029426 Homo sapi
19	1216.5	14.4	926	13 BU146839	BU146839 AGENCOURT
20	1208	14.3	800	13 BU750944	BU750944 CH3#036 H
21	1167	13.8	3890	11 BC053340	BC053340 Mus muscu
22	1152	13.7	769	14 CD349566	CD349566 UT-M-PYO-
23	1147	13.6	756	12 EG435656	EG435656 602506942
24	1107	13.1	724	14 CB962917	CB962917 AGENCOURT
25	1100	13.1	633	14 CA773136	CA773136 im60803.Y
26	1078	12.8	670	12 BG570726	BG570726 602591840
27	1069	12.7	768	13 BU853701	BU853701 AGENCOURT
28	1066	12.7	961	14 CD242938	CD242938 AGENCOURT
29	1063	12.6	713	14 CF745072	CF745072 UT-M-GVO-
30	1058	12.6	623	14 CA773849	CA773849 im58a04.Y
31	1057	12.5	642	14 CA773183	CA773183 im60804.Y
32	1055	12.5	773	13 EQ441353	EQ441353 AGENCOURT
33	1053	12.5	664	14 CB044826	CB044826 NISC_gc06
34	1051	12.5	4585	11 AK089794	AK089794 Mus muscu
35	1049	12.4	627	14 CD767029	CD767029 AGENCOURT
36	1048.5	12.4	772	12 BG173854	BG173854 602333913
37	1036	12.3	614	14 CA771132	CA771132 i071b12.x
38	1024	12.2	687	10 BB341620	BB341620 B3341620
39	1009.5	12.0	3199	11 BC047342	BC047342 Homo sapi
40	993	11.8	658	14 CD351000	CD351000 UT-M-GIO-
41	992	11.8	627	10 BB192788	BB192788 BB192788
42	991	11.8	702	10 BB099251	BB099251 BB099251
43	981.5	11.6	650	10 BB190957	BB190957 BB190957
44	964	11.4	581	14 CA941420	CA941420 i-34b12.x
45	961	11.4	578	13 EU069128	EU069128 im58a04.x

# ALIGNMENTS

RESULT 1  
AK047188  
LOCUS  
DEFINITION  
Mus musculus 10 days neonate cerebellum cDNA, RIKEN full-length  
enriched library, clone:B30033A02 product:ATP-BINDING CASSETTE  
PROTEIN homolog [Homo sapiens], full insert sequence.  
ACCESSION  
AK047188  
VERSION  
AK047188.1 GI:26338625  
KEYWORDS  
HTC; CAP trapper.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
3950 bp mRNA linear HTC 20-SEP-2003

```

Qy 1524 PheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeu----- 1541
Db 6367 TTTGGAGATGGCTATATCGTCACAAATGAGATCAATCCCGAAGGACGACCTGCTTCCT 6426
Qy 1542 GluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGlu 1561
Db 6427 GACCTGAACCTGTGGAGCAGTTCTTCCAGGGGAACTTCCAGGCAGTGTGCAGAGGGAG 6486
Qy 1562 SerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGln 1581
Db 6487 AGGCACATACACATGCTCCAGTTCAGGTCTCCTCC-----TCTTCCCTGGCGAGG 6537
Qy 1582 SerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSer 1601
Db 6538 ATCTTCCAGCTCCTCCTCCCAAGGACAGCCTGCTCATCGAGGAGTACTCAGTCACA 6597
Qy 1602 GlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGlu 1618
Db 6598 CAGACCACACTGGACCGAGTGTGTTGTAATTTTGCTAAACAGCAGACTGAA 6648

```

Search completed: April 13, 2004, 04:19:02  
Job time : 781 secs

Db 4402 GCCGGGGCCCTCCGCCCCAGAGAACACAGCGGACGCGGAATTTCTACAGACCTG 4461  
Qy 921 AlaAspSerAspIleSerAspLeuIleSerPheThrSerGlnAsnIleMetValThr 940  
Db 4462 ACGGACAGGAACATCTCCGACTCTTGGA- 4491  
Qy 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 4492 -----AAACGATCTCTCTCTATTAAGAGACGCTTAAAGAGCAATTTCTCG 4539  
Qy 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 4540 GTCAATGAACAGAGGTATGGAGGAATTTCCATTGGAGAAAGCTCCAGTC-----GTC 4593  
Qy 981 ProIle-----LeuValAsnIleIleSerAsnTyr----- 990  
Db 4594 CCCATCAGGGGGAAGCACTGTGTGGTATTAAGGACCTTGGCCGCGATCATGAATGTG 4653  
Qy 991 -----TyrLeuTyrHisLeu 995  
Db 4654 AGCGGGGCCCTATCACTAGAGAGGCTCTAAAGAAATACCTGATTTCTTAAACATCTA 4713  
Qy 996 AsnValThrGluThrIleGlnIleTTP----- 1004  
Db 4714 GAAACTGAAGACAACATTAAGTGTGTTTAAATACAAAGGCTGGCATCCCTGGTCAAC 4773  
Qy 1004 ----- 1004  
Db 4774 TTCTCAATGTGGCCCAACAGCCATCTTACGGGCGAGCTGTGCTTAAGGACAGGACCCC 4833  
Qy 1005 -----SerThrProPhe-----PheGln 1010  
Db 4834 GAGGAGTATGGAATCACCGTCATTAGCCCAACCCCTGAACCTGACCAAGGACGCTCTCA 4893  
Qy 1011 GluIleThrAspIleValPheLysIleGluLeuTyrPheGlnAlaAlaLeuGlyIle 1030  
Db 4894 GAGATTACAGTGTGACCACTTCAGTGGAT-----GCTGTGGTGGCAATC 4938  
Qy 1031 IleVal-----ThrAlaMetProTyrPheAlaMetGluAsnAlaGluAsn 1046  
Db 4939 TGGGTGATTTCTCCATCTCTCGTCCCAAGCAGCTTTGTCTTTATTTGATCCAGGAG 4998  
Qy 1047 HisIleIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTTP 1066  
Db 4999 CGGGTGAACAAATCCAAAGCACTCCAGTTTATCAGTGGAGTGAGCCCAACCTACTCG 5058  
Qy 1067 IleGlyGlnAlaValAspIleProLeuPhePheIleIleLeuIleLeuMetLeuGly 1086  
Db 5059 GTGACCAACTTCCTCTGGGACATCATGAATATTCCGTGAGTCTGGGCTGGTGGTGGC 5118  
Qy 1087 SerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaVal 1106  
Db 5119 ATCTTCATCGGGTTTCAGAGAAAGCGTACACTTCCAGAAACCTCTCCGCTGGTGG 5178  
Qy 1107 PheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPheThr 1126  
Db 5179 GCATGCTCTGCTGTATGATGATGGGGTTCATTCCTCATGTATACCCAGCA----- 5229  
Qy 1127 PheLysLysIleLeuAsnThrLysGluPheTTPSerPheIleTyrSerValAlaAlaLeu 1146  
Db 5230 -----TCCTTCTGTGTGATGTCCCAACGACA 5256  
Qy 1147 AlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyrThrIleAla----- 1163  
Db 5257 GCCTATGTGCTTATCTGTGTAATCTGTTTCATCGGCATCAACAGCAGTGTATTACC 5316  
Qy 1164 -----ThrIleLeuHisTyr----- 1168  
Db 5317 TTCATCTTGAATATTGTAGATAAACCGACGCTCTCAGGTTCACCGCGTGTGAGG 5376  
Qy 1169 AlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLeuIleSerPheIleLys 1188

Db 5377 AAGCTGCTCATTTGTTCTCCCACTTCTGCTGGCGCGGCGCTCATTTGAC---CTTGCA 5433  
Qy 1189 IleserTrpLysAsnValArgLysAsnValAspThrTyr----- 1201  
Db 5434 CTGAGC-----CAGGCTGTGACAGATGCTATCCCGGTTTGTGTGAGGACAC 5481  
Qy 1202 -----AsnPro-----TrpAspArgLeuSerValAlaValIleSerProTyrLeuGln 1217  
Db 5482 TCTGCAATCCGTTCCACTGGACCTGATTGGGAAGAACCTGTGTGCCATGCTGGTGA 5541  
Qy 1218 CysValLeuTrrPheLeuLeuGlnTyrTyrGluLysLysTyrGlyGlyArgSerIle 1237  
Db 5542 GGGGTGGTGTAC---TTCTCTCTCACTGCTGCTGCTCCAGCGCCAC----- 5593  
Qy 1238 ArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGlu 1257  
Db 5594 -----TTCTTC-----CTCTCCCAATGGATTGCCGAG 5610  
Qy 1258 ProProAspAsnGlu-----AspGluAspGluAspValLysAlaGluArgLeuLysVal 1275  
Db 5611 CCCACTAAGGAGCCCATTTGTTGATGAAGATGATGCTGAGTGAAGAAAGCAAGAATT 5670  
Qy 1276 LysGluLeuMetGlyCysGlnCysCysGluLysProSerIleMetValSerAsnLeu 1295  
Db 5671 -----ATTACTGCT----- 5679  
Qy 1296 HisLysGluTyrAspAspLysLysAspPheLeuLeuSerArgLysValLysVal--- 1314  
Db 5680 -----GGAAATAAACTGACATCTTAAGGCTACATGAACCTAAACAAGATTAT 5727  
Qy 1315 -----AlaThrLysTyrIleSerPheCysValLysLysGlyGluIle 1328  
Db 5728 CCAGGCCTCCAGCCAGCAGTGGACAGGCTGTGTGCGAGTTCGCCCTGGAGATGC 5787  
Qy 1329 LeuGlyLeuLeuGlyProAsnGlyValagLysSerThrIleIleAsnIleLeuValGly 1348  
Db 5788 TTTGGCTCTCTGGGAGTGAATGTGTGCGGCAAAACCAACCATTCATCAAGATGCTCACTGGG 5847  
Qy 1349 AspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSer----- 1363  
Db 5848 GACACCAAGTGCCTCAGGGGATGCCACGTCAGCGCAAGAGTATTTTAAACCAATTT 5907  
Qy 1364 SerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsn 1383  
Db 5908 TCTGAGTCCATCAAAAT-----ATGGCTACTGTCTCTCAGTTGAT 5949  
Qy 1384 ProLeuTTPProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGly 1403  
Db 5950 GCAATCGATGAGCTGCTCACAGAGCAGAACATCTTTACCTTTATGCCCGCTTCGAGGT 6009  
Qy 1404 MetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLys 1423  
Db 6010 GTACAGCAGAGAAATCGAAAGTTGCAAACTGGAGTATTAGAGCTGGGCTGACT 6069  
Qy 1424 GluHisLeuGlnLysThrValLysLeuProAlaGlyIleLysArgLysLeuCysPhe 1443  
Db 6070 GTCTACGCGGACTGCTGGCTGGCAGCGTACAGTGGGGGCAACAAGCGGAAACTCTCCACA 6129  
Qy 1444 AlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMet 1463  
Db 6130 GCCATCGCATCTATGGCTGCCCGCTGCTGCTGTGATGAGCCACCAAGGATG 6189  
Qy 1464 AspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLys 1483  
Db 6190 GACCCCAAGGACGCGCATGCTGTGGAACGTCATCGTGCAGCATCATCAGA---GAAGGG 6246  
Qy 1484 ArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgVal 1503  
Db 6247 AGGGCTGGTGTCTCATCATCCAGCATGGAAATGTGAGGCACCTGTGTACCCGCTG 6306  
Qy 1504 AlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLys 1523  
Db 6307 GCCATCATGTGTAAAGGGCGCTTTTCGATGTATGSGCACCATTTCAGCATCTCAAGTCCAAA 6366

QY 335 -----ThrValAlaPheGlyPheIleGlyLeuMetIleLeuIleLeuSer 350  
Db 2317 TACCTGCCACACATCCTCTGCTTCCGCTGCAGGACCGCATGACGGTGAGCTG----- 2370  
QY 351 PheProLysSerLeuValTrpLeuPheSerProPheCysHisCysThrPheValIleGly 370  
Db 2371 -----AAGAAGGCTGTGAGCTTACTGCTCCG-----GTGGCAITTGATTGGC 2415  
QY 371 IleAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSerPheSerAsnLeuThr 390  
Db 2416 ACTGAGTACCTGGTTCGCTTTGAAGAGCAAGCGCTGGGGCTGCAGTGGAGCAACATCGGG 2475  
QY 391 AlaGlyPro-----TyrProLeuIleIleThrIleLeuMetLeuThrLeu 405  
Db 2476 AACAGTCCCACGGAAGGAGCGAATTCAGCTTCTGCTGCTCCATGCAGATGATGCTCTT 2535  
QY 406 AsnSerIlePheTrpValLeuLeuAlaValTyrLeuAspGlnValIleProGlyGluPhe 425  
Db 2536 GATGCTGCTGCTATGCTTACTGCTTGTGATGCTTGTGATGCTTGTGATGCTTGTGATGCT 2595  
QY 426 GlyLeuArgArgSerLeuTyrPheLeuLysProSerTyrTrp----- 440  
Db 2596 GGAACCCCACTCTCTGCTGCTTCTTCTACAGAGTCGTATTGGCTGGCGGTGAAGG 2655  
QY 441 -----SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnVal 455  
Db 2656 TGTTCAACACAGAGAAGAGCGCTCGAAAAGACCGGCCCTTAACAGAGGAACGGAG 2715  
QY 456 AsnGlyAsnIleSerPheSerGluIleIleGluProValSerSerGluPheValGlyLys 475  
Db 2716 GAT-----CGAGCACCCAGAGGAATAACAGACTCTCTTCTTTGACGT 2760  
QY 476 Glu-----AlaIleArgIleSerGlyIleGlnLysThrTyrArgLys 489  
Db 2761 GAGCATCCAGGCTGGTTCCTGGGTATGCTGTGAAGAATCTGTGAAGAATTTTGTAGCCC 2820  
QY 490 LysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIle 509  
Db 2821 TGTGGCCGG-----CGAGCTGTGGACCGTCTGAACATACCTTCTACGAGAACCCAGATC 2874  
QY 510 ThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGly 529  
Db 2875 ACCGATTCCTGGGCCCAATAGGAGCTGGGAAACCCACCACTTGTCTCATCTCTGACGGT 2934  
QY 530 LeuCysProProSerAspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAsp 549  
Db 2935 CTGTGCCACCAACCTCTGGGACTGTGCTCTGCTGGGGGAAG-----GACATTGAA 2985  
QY 550 GluMetPheGluAla---ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPhe 568  
Db 2986 ACCAGCCTGGATGCAGTCCGCGCAGAGCTTGGCATGTGTCCACAGCAACATCCTGTTT 3045  
QY 569 AspValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAla 588  
Db 3046 CACCACCTCAGCGTGGCTGACATCATCTGTATCCAGCTGGAAGAAAGTCCAG 3105  
QY 589 AsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuMetGlnThrIleLys 608  
Db 3106 GAGGAGCCAGCTGGAGATGGAGAGCCATGTTGGAGGACAGAGCTCCACCAAGCGG 3165  
QY 609 AspAsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAla 628  
Db 3166 AATGAAGAGGCTCAGGACCTATCAGTGGCATGACAGAGAAAGTGTGCGTGGCATTTGCC 3225  
QY 629 ValLeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCys 648  
Db 3226 TTTGGGAGATGCAAGGTGGTGTATCTGAGCAACCACTCTGGGGTGGACCTTAC 3285  
QY 649 SerArgHisIleValTrpAsnLeuLysTyrArgLysAlaAsnArgValThrValPhe 668  
Db 3286 TCGAGACGCTCAATCTGGGATCTGCTCTGAAGTATCGCTCAGGCAAGAACCATCATCATG 3345

QY 669 SerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGln 688  
Db 3346 TCCACTCACCACTGACGAGCGGACCTCTCTGGGACCGCATGTCATTCATTCGCCAG 3405  
QY 689 GlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyr 708  
Db 3406 GGAAGGCTCTACTGCTCAGGACCGCCACTCTCTCTGAAGAACTGCTTTGGCACAGGCTT 3465  
QY 709 ArgLeuSerMet----- 712  
Db 3466 TACTTAACCTTGGTGGCGCAAGATGAAAACATCCAGAGCCAAAGAAAGGAGGTGAGGG 3525  
QY 713 -----TyrIleAspLysTyr 717  
Db 3526 ACCTGCAGCTGCTGCTCTAAGGGTTTCTCCACACAGTGTCCAGCGCCAGTGCATGACCTA 3585  
QY 718 CysAlaThrGlu-----SerLeuSerSerLeuValIleGlnHis 730  
Db 3586 ACTCCAGAACCAAGTCTCGATGGGATGTAATAGCTGATGATGATGATGATGATGATGAT 3645  
QY 731 IleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPhe 750  
Db 3646 GTTCCAGAGGCAAGAGCTGCTGAGTGCAATTGCTCAAGAACTATCTCTCTCTTCCAAT 3705  
QY 751 LysAspMetAsp-----LysPheSerClyLeuPheSerAlaLeu---AspSerHisSer 767  
Db 3706 AAGAATCTCAAGCACAGAGCATATCCAGCCTTTTCAGAGAGCTGGAGAGAGCTGGCT 3765  
QY 768 AsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLys 787  
Db 3766 GACCTTGTCTCAGCAGTTTGGAAATTTCTGACACTCCCTCTGGAAGAGATTTTCTGAAG 3825  
QY 788 LeuGluValGluAlaGluIle----- 794  
Db 3826 GTCACGGAGGATTTCTGATTACGAGACCTCTGTTTGGGGTGGCTGCTACAGAAAGAGAA 3885  
QY 795 -----AspGlnAla 797  
Db 3886 AACGTCAACCCCGACACCCCTGCTTGGTCTCCAGAGAGAGGTGGACACACACCCAG 3945  
QY 798 AspyrSerValPheThr-----GlnGlnProLeu 807  
Db 3946 GACTCCAATGTCTGCTCCCGGCGCGCTGCTCACCAGAGGCGCAGCTCCGCCCA 4005  
QY 808 GluGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuIleLeuSer 827  
Db 4006 GAGCCAGAGTCCCGAGCCCGCAGCTCAACAGGGGACACAG-----CTGGTCTCCAG 4059  
QY 828 GluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGlnMetTyrThrIle 847  
Db 4060 CATGTGCAGCGCTGCTGCTC-----AAGAGATTCCAACACACCATC 4101  
QY 848 AlaLysPheHisPhePheThrLeuLysArgGluSerLysSerValArgSerValLeuLeu 867  
Db 4102 CGACGCCACAGGACTTCTGGCGCAGATCGTCTCCCGGTACCTTGTGTTGTTTGGCT 4161  
QY 868 LeuLeuLeuIlePhe-----PheThrValGlnIle 877  
Db 4162 CTGATGCTTCTATTGTTATCCTCTTTCGGCAATACCCCGCTTTTGACCTTCACCCC 4221  
QY 878 PheMetPheLeuValHisSerPheLysAsnAlaValPro-----Ile 893  
Db 4222 TGGATATATGGCAGCAGTACACCTTCTCAGCATGGATGAACACGAGCGAGTGAGCATTC 4281  
QY 894 LysLeuValProAspLeuTyrPheLeuLysProGly----- 905  
Db 4282 ACGGTACTTGCAGAGTCTCTCTGAATAAGCCAGGCTTTGGCAACCGCTGCTGAGGAA 4341  
QY 905 ----- 905  
Db 4342 GGGTGGCTTCCGTGCAGCACCAGGAGAGCTCACCATGCTGCCAGAGTGGCCGAGGT 4401  
QY 906 -----AspLysProHisLysTyrLysThrSerLeu---LeuLeuGlnAsnSer 920

```
QY 1423 LysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuGluCys 1442
Db 6416 GTCAAGTATCGAAGAAATATGCTGTAATACTAGTGGAGGCAACAAAGCAAGCTCTCT 6475
QY 1443 PheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGly 1462
Db 6476 ACAGCATGCTCTTGTATCGGCGGCTCTCTGTTCTGTGATGAACCCACACAGGC 6535
QY 1463 MetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLysAsnArg 1482
Db 6536 ATGATCCCAACACCGCGGTTCTTGTGAATTGTGCCCTAAGTGTCTCAAG---GAG 6592
QY 1483 LysArgAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCysAspArg 1502
Db 6593 GGGAGATCATAGTGTCTATCATCTAGTATGGAAGAATGTGAAGCTCTTTGCACATGG 6652
QY 1503 ValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSer 1522
Db 6653 ATGGCAATCATGCTCAATGAAGGTCAGTGCGCTTGGCAGTGTCCAGCATCTAAANAAT 6712
QY 1523 LysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeuGlu 1542
Db 6713 AGTTTGGAGATGTTATACATATGATGTACGAATAGCAGGG-----TCCAACCCVRGAC 6766
QY 1543 ValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSer 1562
Db 6767 CTGAAGCCTGTCAGGATTTCTTTGGACTTGGCATTTCTTGGAGTGTCTTAAGAGAGAAA 6826
QY 1563 PheSerSerIleLeuAlaTyrIleProLysGluAspValGlnSerLeuSerGlnSer 1582
Db 6827 CACCGGAACATGCTACATACACAGCTTCCA---TCTTCATTATCTTCTCTGCCAGGATA 6883
QY 1583 PhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGln 1602
Db 6884 TTCAGCATCTCTCCACAGCAAAAGCGACTCCACATAGAAGACTACTCTGTTCTCAG 6943
QY 1603 AlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAspAsn--- 1621
Db 6944 ACACACTTGACCAAGTATTTGGAAGTGTGCAAGGACCAAGATGATGATGACACATTA 7003
QY 1622 -----SerCysGlyThrLeuAsnSerThr 1629
Db 7004 AAAGACCTCTCATTAACAAAAACCAGACAGTAGTGGANMTTGCAGTCTCTCATCTTTT 7063
QY 1630 LeuTyrTrpGluArgThrGlnGluAsp-----ArgValValPhe 1642
Db 7064 CTACAGATGAGAAAGTGAAGAAAGCTATGATGAGAAATCTCTGTTTC 7111
```

## RESULT 15

```
US-09-032-438C-5
; Sequence 5, Application US/09032438C
; GENERAL INFORMATION:
; APPLICANT: Ratner, Amir
; APPLICANT: Sun, Hui
; APPLICANT: Lupski, James R.
; APPLICANT: Nathans, Jeremy
; APPLICANT: Anderson, Kent L.
; APPLICANT: Leppert, Mark
; APPLICANT: Dean, Michael
; APPLICANT: Singh, Nanda
; APPLICANT: Shroyer, Noah F.
; APPLICANT: Smallwood, Philip M.
; APPLICANT: Allikmets, Rando
; APPLICANT: Lewis, Richard A.
; APPLICANT: Li, Yixin
; TITLE OF INVENTION: Nucleic Acid And Amino Acid Sequences For ATP-Binding Cassette
; TITLE OF INVENTION: Transporter And Methods Of Screening For Agents That Modify
; TITLE OF INVENTION: ATP-Binding Cassette Transporter
; FILE REFERENCE: BYLR-0065
; CURRENT APPLICATION NUMBER: US/09/032,438C
; CURRENT FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: US 60/039,388
; PRIOR FILING DATE: 1997-02-27
```

```
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 6705
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-032-438C-5

Alignment Scores:
Pred. No.: 5,53e-123 Length: 6705
Score: 1363.00 Matches: 458
Percent Similarity: 42.62% Conservative: 325
Best Local Similarity: 24.93% Mismatches: 630
Query Match: 16.18% Indels: 424
DB: 5 Gaps: 59
```

US-10-090-458-5 (1-1642) x US-09-032-438C-5 (1-6705)

```
QY 84 ValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeuProAspVal 103
Db 1504 ATATTTAACTACATCATGATCGACCCCTCCGCTGTGTCAT---CAATACCTGGAGTGTCTG 1560
QY 104 IleIleThr-----GluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeuSer 121
Db 1561 GTCCTGGATAAGTTTGAAGCTACATGATAAAGTCAAGCTCAGCTCACCACCAAGTGCCTCTCT 1620
QY 122 -----LysProSerAsnPheValGlyValValPheLysAsp----- 133
Db 1621 CTACTGGAGGAAACAATGTTCTGGCGCGAGTGTATTCCTCGCATGTATCCCTGGACC 1680
QY 134 -----SerMetSerTyrGluLeuArgPhePheProAspMetIle----- 146
Db 1681 AGCTCTTACACCCACCGAGTGAAGTATAGATCGAATGCAATGACATAGACGTGGTGAGAAA 1740
QY 147 -----ProValSerSerIleTyrMetAspSer-----ArgAlaGlyCysSerLysSer 162
Db 1741 ACCAATAAGATTAAGACACAGGTATGGGATTCCTGCCAGAGCT-----GATCCC 1791
QY 163 CysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIleAspAla 182
Db 1792 GTGAAGATTTCCGGTACATCTGGGCGGGTTCCTATCTGCAGGACNTGGTGTGAACAG 1851
QY 183 AlaIle-----IleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGlu 198
Db 1852 GGGATCACAGAGCCAGGTGCGAGCGAGGCTCCAGTTCGAATCTACCTCCAGCAGATG 1911
QY 199 SerThrLysAlaValIleMetGlyGluThrAlaValValGluLeuAspThrPheProArg 218
Db 1912 CCTTACCCCTGCTTGGTGGAGCATTTCTTTCATGATCATCTCGAACCCGCTGTTTCCCT--- 1968
QY 219 GlyValIleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIle 238
Db 1969 -----ATCTTTCATGGTGTGTCATGATCTACTCTGTCTCTCCATGACTGTGAAG 2016
QY 239 HisIleValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHis 258
Db 2017 AGCATCGTCTTGGAGAAGAGTTCGCACTGAAGAGACCTTGAAATAATCAGGGTGTCTCC 2076
QY 259 AspThrAlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSer 278
Db 2077 AATCAGTGTATTTGGTGTACTTCTGTCAGCAGCTTCTCCATCATGTCTGATGAGATC 2136
QY 279 LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleVal 298
Db 2137 TTCCTCTCGCATATTTCATCATGTCATGGAAGATCTCTACATTACAGCAGCCCATTCATC 2196
QY 299 IlePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThr 318
Db 2197 CTCTTCTGTTCTTGTGGTCTTCTCCACTGCCACCATCATGCTGTGCTTCTGCTCAGC 2256
QY 319 ProLeuPheLysLysSerLysHisVal-----GlyIleValGluPhePheVal 334
Db 2257 ACCTTCTTCTCCAGGCGAGTGGCAGCAGCTGTAGTGTGTGTCTCATCTATTTCACCCCTC 2316
```

Db 4415 -----GTGCCACCCCTTTGGCAAGTACCCCGAGCTCGAAGCTTCAG 4453  
 Qy 900 -----TyrPheLeuLysProGlyAspLysProHisLysTyrLys 912  
 Db 4454 CCCTGGATGTACAACGAACAGTACACATTTGTGAGCAATGATGCTCTGAGGACACGGGA 4513  
 Qy 913 ThrSerLeuLeuLeu----- 917  
 Db 4514 ACCCTGGAATCTTTAAAGCCCTCACCAAGACCCCTGGCTTCGGGACCCGCTGTATGGAA 4573  
 Qy 917 ----- 917  
 Db 4574 GGAACCAATCCAGACACRCCTCGCCAGGAGGAGGAGAGTGGACCACTGCCCA 4633  
 Qy 917 ----- 917  
 Db 4634 GTTCCCGAGACCATCATGAGACCTTTCAGATGGGAACCTGGCAATGAGAACCTTCA 4693  
 Qy 918 -----GlnAsnSerAlaAsp----- 922  
 Db 4694 CCTGCATGCCAGTGTAGCAGCGACAAATCAAGAAGATGCTGCTGTGTCGCCCGAGK 4753  
 Qy 923 -----SerAspIleSerAspLeuIleSerPhePhe 932  
 Db 4754 GCAGGGGGTGCCTCTCCACAAAGAAACAAACACACTGCAGATATCTTCAGGACCTG 4813  
 Qy 933 ThrSerGlnAsnIleMetValThrMet----- 941  
 Db 4814 ACAGGAAGAACATTTTGGATTATCTGCTGAGACGATATGTCAGATCATAGCAAAAGC 4873  
 Qy 941 ----- 941  
 Db 4874 TTAAGAACAAGATCTGGGTGAATGAGTTTGGTATGAGCGCTTTCCCTGGGTGTCAGT 4933  
 Qy 942 -----IleAsnAspSer----- 946  
 Db 4934 AATACTCAAGCACTTCTCCAGTCAAGAAGTAAATGATCCCAACAAATGAAGAA 4993  
 Qy 947 TyrValSerValAlaProHisSerAlaAlaLeuAsnValMethHisSerGluLysAspTyr 966  
 Db 4994 CACCTAAAGCTGGCCAGGACAGTCTGCGAGATCGATTTCTCAACAGCTTGGGAAGATT 5053  
 Qy 967 Val-----PheAlaAlaValPheAsnSerThrMetValTyr 978  
 Db 5054 ATGACGAGCTGGACACCAAAATATGTCAAGTGTGTTCAATAACAGGCTGGCAT 5113  
 Qy 979 SerLeuProIleLeuValAsnIle----- 987  
 Db 5114 GCAATCAGCTTTTCTCGTAATGTATCAACRATGCCATTCCTCCGGGCCACCTGCAAAAG 5173  
 Qy 988 -----SerAsnTyr-----TyrLeuTyrHisLeuAsnValThrGlu 999  
 Db 5174 GGAGAGAACCTAGCCATTAAGAAATGATCTGCTTTCAATCATCCCTGAATCTCAACAG 5233  
 Qy 1000 ThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIle 1019  
 Db 5234 CAG---CAGCTCTCAGAGGTGGCTCYGATGACCAATCATGATGATGCTTGTGTCATC 5290  
 Qy 1020 GluLeuTyrPheGlnAlaAlaLeuLeuGlyLleIleValThrAlaMetProProTyrPhe 1039  
 Db 5291 TGTGTATCTTT-----GCAATGTCCTTGTGCCCGCCAGCTTT 5329  
 Qy 1040 AlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGly 1059  
 Db 5330 GTCGTATTCTGATCCAGGAGCGGTGAGCAAAAGCAAAACACCTGCGAGTTCATCAGTGA 5389  
 Qy 1060 LeuLeuProSerAlaTyrTrpIleGlyGluAlaValValAspIleProLeuPhePheIle 1079  
 Db 5390 GTGAAGCTGTCACTACTGCTCTCAATTTTGTCTGGGATATGTGCAATACGTTGTC 5449  
 Qy 1080 IleLeuIleLeuMetLeuLeuLysSerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThr 1099  
 Db 5450 CCTGCCACACTGGTCAATTATCATCTTCTGCTTCCAGCAGAGTCTCATGTGCTCTCC 5509

Qy 1100 ValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPhe 1119  
 Db 5510 ACCAATCGCTGCTGATGACCTTCTACTTTTGTGTATGGGTGCAATCAACACCTCTC 5569  
 Qy 1120 ThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPhe 1139  
 Db 5570 ATGTACCCAGCTCTTTGTGTTCAAG----- 5596  
 Qy 1140 IleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGly 1159  
 Db 5597 -----ATCCCCAGCAGACCTATGCTGTGCTCCAGCGTGAACCTCTTCTATGGC 5647  
 Qy 1160 Tyr----- 1161  
 Db 5648 ATTAATGCACGCTGCCACCTTTGTGCTGAGCTGTTTCACYGACAATAAGCTGAATAAT 5707  
 Qy 1162 IleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleThrProLeuLeuGly 1181  
 Db 5708 ATCMATGATATCTCAAGTCCGTGTTT-----TTGATCTTCCACATTTTTCCTGGGACGA 5764  
 Qy 1182 CysLeuIleSerPheIleLys----- 1188  
 Db 5765 GGGCTCATCGATCGTGGAAACACGCAATGCTGATGCCCTGGAAAGTTTGGGAG 5824  
 Qy 1189 -----IleSerTrpLysAsnValArgLysAsnValAspThrTyrAsn 1202  
 Db 5825 AATCGCTTTGTGTCACCATATCTTGGGACTTGTGGGAYGAAACCTCTTCGCC- 5878  
 Qy 1203 ProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIle 1222  
 Db 5879 -----ATGGCGGTGGAAGGGTGTGTTCTTCTCTTACTACTGTT- 5917  
 Qy 1223 PheLeuLeuGlnTyrTyrGluLysTyrGlyArgSerIleArgLysAspProPhe 1242  
 Db 5918 ---CTGATCCAGTAC-----AGATTC 5935  
 Qy 1243 PheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGlu 1262  
 Db 5936 TTC-----ATCAGCCCGACACCTGTAAATGCAAGCTA-----TCTCTCTGAAAT- 5980  
 Qy 1263 AspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGln 1282  
 Db 5981 GATGAAGATGAAGATGTGAGCGGGAAGACAGAGAATCTTGTATGCTGGAGGC- 6034  
 Qy 1283 CysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLys 1302  
 Db 6035 -----CAGAATGACATCTTAGAAATCAAGAGTGTGCAAGATATAT- 6076  
 Qy 1303 LysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCys 1322  
 Db 6077 -----AGAAGGAAGCGGAAGCTGCTGTTGACAGATTTGCTGGGCGC 6118  
 Qy 1323 ValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThrIle 1342  
 Db 6119 ATTCTCTCTGTGATGCTTTGGGCTCTGAGGATTAATGGGCTGGAATCATCAACT 6178  
 Qy 1343 IleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyr 1362  
 Db 6179 TTCAGATGTTTAAACAGAGATACCATGTTTACCAGAGAGATGCTTTCCTT- 6235  
 Qy 1363 SerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIle 1382  
 Db 6236 AATAGTATCTTATCAACATCATCAAGATACATCAACATGGGCTACTGCCCTCAGTTT 6295  
 Qy 1383 AsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLys 1402  
 Db 6296 GATGCCATCAGAGCTGTTGCTGGGAGAGAACACGTGGAGTCTTGTGCCCTTTTGAGA 6355  
 Qy 1403 GlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeu 1422  
 Db 6356 GGAGTCCAGAGAAAGATTGGCAAGGTTGTTGAGTGGGCGGATTCGGAACCTGGGCTC 6415



Qy 262 PheTrpLeuSerTrpValLeuLeuTyrrThrSerLeuIlePheLeuMet-----Ser 278  
Db 2354 CTCGGTTTACGTGGTTCATT-----AGTAGCCTCATCTCTCTTGTGAGCGGTGGC 2407  
Qy 279 LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleVal 298  
Db 2408 CTCGTAGTGTCTCCCTGAGATAGGAACCTGCTG---CCCTACAGTGCATCCAGCGTG 2464  
Qy 299 IlePheLeuLeuPhePheLeuTyrrGlyLeuSerSerValPhePheAlaLeuMetLeuThr 318  
Db 2465 GTGTTTCTCTCTGTCCTCGGTGTTTGTGCTGTGTGTGACATCTCTGCAGTGCCTTCTCGATTAGC 2524  
Qy 319 ProLeuPheLysLysSerLysHisVal-----GlyIleValGluPhePheVal 334  
Db 2525 ACATCTCTCCAGAGCCAACCTGGCAGCAGCCTGTGGGGGCATCATCTACTTACAGCTG 2584  
Qy 335 ThrValAlaPhe-----GlyPheIleGlyLeuMetIleLeuLeu 347  
Db 2585 TACCTGCCCTACGTCCTGTGTGGCATGGCAGCACTACTGTGGCTTCMACCTCAAGATC 2644  
Qy 348 IleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCysThrPhe 367  
Db 2645 TTGCTAGC-----CTGCTGTCTCT-----GTGGCTTTT 2674  
Qy 368 ValIleGlyLeAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSerPheSer 387  
Db 2675 GGGTTTGGCTGTAGTACTTTTGGCCCTTTTGGAGCAGCGGCATTTGGAGTGCAGTGGAG 2734  
Qy 388 AsnLeuThrAlaGlyPro-----TyrProLeuIleIleThrIleMet 402  
Db 2735 AACCTGTTTGAGAGTCTGTGGAGGAAGATGGCTTCAATCTCACCACTTCGRCTCCATG 2794  
Qy 403 LeuThrLeuAsnSerIlePheTrpValLeuAlaValTrpLeuAspGlnValIlePro 422  
Db 2795 ATGCTGTTTRACACCTTCTCTATGAGGTGTGATGACCTGTGTACATTGAGGCTGTCTTCCA 2854  
Qy 423 GlyLeuPheGlyLeuArgSerSerLeuTyrrPheLeuLysProSerTyrrTrp----- 440  
Db 2855 GCCCAGTACGGAATTCACAGCCCTCGTATTTTCTCTGACCAAGTCTCTACTGGTTGGC 2914  
Qy 441 -----SerLysSerLysArgAsnTyrrGluGluLeuSerGluGlyAsn 454  
Db 2915 RAGGAAGATGATGAGAAGAGCCACCTGTGTTCCAACAGAGAAGATTCAGAA----- 2968  
Qy 455 ValAsnGlyAsnIleSerPheSerGluIleGluProValSerSerGluPheValGly 474  
Db 2969 -----ATCTGCATGGAGGAG-----GAACCCACCCACTTG----- 2998  
Qy 475 LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrrArgLysLysGlyGluAsnVal 494  
Db 2999 AAGCTGGCGTGTCCATTCAGAACCTGGTAAAGTCTAC---YGATGGGATGAAGGTG 3055  
Qy 495 GluAlaLeuArgAsnLeuSerPheAspIleTyrrGluGlyGlnIleThrAlaLeuLeuGly 514  
Db 3056 ---GCTGTCGATGGCTGGCACTGAATTTTATGAGGGCCAGATCACTCTCTCTCTGGG 3112  
Qy 515 HiserGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysProProSer 534  
Db 3113 CACATGGAGTGGGAGACACACCATGTCATCTGACCGGTTGTTTCCCGCCGACC 3172  
Qy 535 AspGlyPheAlaSerIleTyrrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
Db 3173 TCGGGCACYGCATACCTCTGGGAAAAAGACATTGCG-----TCTGAGATGAGCACCATC 3226  
Qy 555 ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
Db 3227 CGGCAACCTGGGGGTCTGTCCCGACATACAGTGTCTGTGTGTACATGCTGACTGCGAA 3286  
Qy 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGlu 594  
Db 3287 GAACACATCTGGTTCTATGCCCCGCTTGAAAGGCTCTCTRAGAAGCACCTGAAGCGGAG 3346  
Qy 595 ValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLys---AspAsnGlnAlaLys 613

Db 3347 ATGGAGCAGATGGCCCTGGATGTTGGTTGGCATCAAGCAGCTGAAAGCAAAACAGC 3406  
Qy 614 LysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuLeuGlyAsnPro 633  
Db 3407 CAGCTGTGAGTGAATGCAGAGAAAGCTATCTGTGGCTTGGCTTTGGCGGGATCT 3466  
Qy 634 LysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleVal 653  
Db 3467 AAGGTGTCATCTTGGATGAACCCACAGCTGTGTGGACCTTACTCTCCGAGGGGAATA 3526  
Qy 654 TrpAsn---LeuLeuLysTyrrArgLysAlaAsnArgValThrValPheSerThrHisPhe 672  
Db 3527 TGGGAGCTGCTGTGAATACCGACAA---GGCGCACCATATTCTCTCTACACACCAC 3583  
Qy 673 MetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLys 692  
Db 3584 ATGGATGAAGCGGACGCTCTGGGKACAGGATGCCATCATCTCCATGGGAAGCTGTGC 3643  
Qy 693 CysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrrArgLeuSerMet 712  
Db 3644 TGTGTGGCTCTCTCCCTGTTCTGAAGAACCAGTGGGAACAGGCTACTACCTGACCTTG 3703  
Qy 712 ----- 712  
Db 3704 GTCAAGAAAGATGGGAATCCTCCTCAGTTCCTGCAGAAACAGTAGTACCTGTGTCA 3763  
Qy 713 TyrrIleAspLysTyrrCysAla----- 719  
Db 3764 TACCTGAAAAAGGAGGACAGTGTCTCAGAGCAGTCTGTGATGTGGCTGGGCGAGCAG 3823  
Qy 720 -----ThrGluSerLeuSerSerLeuValLysGlnHis 730  
Db 3824 CATGASAGTGCACGCTGACCATCGATGCTCTCTGTATCTCCAACTCATCAGGAAGCAT 3883  
Qy 731 IleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrrSerLeuProPhe 750  
Db 3884 GTGCTGAAGCCCGCTGTGGAAGACATAGGGCATGAGCTGACCTATGTCTGCCATAT 3943  
Qy 751 LysAspMetAspLys-----PheSerGlyLeuPheSerAlaLeuAspSerHis---Ser 767  
Db 3944 GARGCTGTAAGGAGGAGCGCTTTGTGAACTTTTCATGAGATTGATGACCGCTCTCA 4003  
Qy 768 AsnLeuGlyValIleSerTyrrGlyValSerMetThrThrLeuGluAspValPheLeuLys 787  
Db 4004 GACCTGGGCATTTCTAGTTATGGCATCTCAGAGACGACCCCTGGAAGAATAATTCTCAAG 4063  
Qy 788 LeuGluValGluAlaGluIleAsp----- 795  
Db 4064 GTGGCCGAAGAGAGTGGGGTGATGTCAGACCTCAGATGCTACCTTGGCCAGCAAGACA 4123  
Qy 796 -----GlnAlaAspTyrrSerValPheThrGln----- 804  
Db 4124 AACAGCGCGGCTTCCGGGACACAGAGCTGTCTTCCCGCTTCACTGAAGATGATGCT 4183  
Qy 805 ---GlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu 823  
Db 4184 GCTGATCCAAATGATCTGCACATAGACCAGATCC-----AGAGAGACAGACTTG 4234  
Qy 824 LeuLeuLeuSerGluThrLysAlaSerLeu---ValSerThrMetSerLeuTrpLysGln 842  
Db 4235 CTCAGTGGGATGGATGGCAAGGRTCTCTACAGGTGAAAGGCTGGAACTTACACAGCAA 4294  
Qy 843 GlnMetTyrrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLysSerVal 862  
Db 4295 CAGTTTGGGCCCTTTTGTGGAAGAGACTGCTAATTGCCAGACGAGTGGGAAGGATTT 4354  
Qy 863 ArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPheLeuVal 882  
Db 4355 TTTGCTCAGATTGCTTGGCCAGCTGTGTTGCTGTCATGGCCCTGTGTTGTTCAGCCTGATC 4414  
Qy 883 HisHiserPheLysAsnAlaValValProIleLysLeuValProAspLeu----- 899

Db 5823 GATGAAGATGAGATGAGCGGGAAGACAGACAGAAATTCCTGATGGTGAGGC----- 5876  
 Qy 1283 CysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLys 1302  
 Db 5877 -----CAGATGACATCTTAGAAATCAAGGAGTTGACGAAGATATAY----- 5918  
 Qy 1303 LysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCys 1322  
 Db 5919 -----AGAGAGAGCGGAAGCGCTGCTTGCACAGGATTCGCTGGGC 5960  
 Qy 1323 ValLysLysGlyGluIleLeuGlyLeuGlyProAsnGlyProAsnGlyLysSerThrIle 1342  
 Db 5961 ATTCCTCTGCTGAGTCTTGGGCTCTGGGAGTTAATGGGCTGGGAAATCATCAACT 6020  
 Qy 1343 IleAsnIleLeuValGlyAspIleLupProThrSerGlyGlnValPheLeuGlyAspTyr 1362  
 Db 6021 TTCAGATGTTACAGAGATACACCTGTACACAGAGAGATGCTTTCCTT---ACAAA 6077  
 Qy 1363 SerSerGluThrSerGluAspAspSerLysCysMetGlyTyrCysProGlnIle 1382  
 Db 6078 AATAGTATCTTATCAACATCCATGAAGTATACATCAGAACATGGGCTACTGCCCTCAGTTT 6137  
 Qy 1383 AsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLys 1402  
 Db 6138 GATCCATCACAGAGCTGTGACTGGGAGAGAACACGTGAGTTCCTGGCCCTTTTGAGA 6197  
 Qy 1403 GlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeu 1422  
 Db 6198 GGAGTCCACAGAAAGAAAGTGGCAAGTTGGTGGTGGCGCATTCGGAACTGGGCGCTC 6257  
 Qy 1423 LysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCys 1442  
 Db 6258 GTGAAGTATGAGAAATAATGCTGGTAATAGTGGAGGCAACAAACGCAAGCTCTCT 6317  
 Qy 1443 PheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGly 1462  
 Db 6318 ACAGCATGGCTTGTGATCGGCGGCTCTGTGGTGTTCCTGGATGAACCCACACAGGC 6377  
 Qy 1463 MetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLysAsnArg 1482  
 Db 6378 ATGATCCCAAGCCGCGGCTCTTGTGAATGTGCCATGTGCTCAAG---GAG 6434  
 Qy 1483 LysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCysAspArg 1502  
 Db 6435 GGGAGATCAGTAGTCTTACATCTCATAGTATGGAAGATGTGAAGCTCTTGCACATAGG 6494  
 Qy 1503 ValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSer 1522  
 Db 6495 ATGCAATCATGTCATGGAAGTTCAGGTGCTTGGCAGTGTCCAGCATCTTAAATAAT 6554  
 Qy 1523 LysPheGlyLysGlyTyrPheLeuGluIleLysLysLysAspTyrIleGluAsnLeuGlu 1542  
 Db 6555 AGGTTGGAGATGGTTATACATAGTTGTAGCAATACAGGG-----TCAAACCTRGAC 6608  
 Qy 1543 ValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSer 1562  
 Db 6609 CTGAAGCCTGTCCAGGATTTCTTTGGACTTGCATTTCTGGAATGTCTTAAAGAGAAA 6668  
 Qy 1563 PheSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSer 1582  
 Db 6669 CACCGAATGCTACAAATACCACTTCCA---TCTTCATATCTTCTTGGCCAGATA 6725  
 Qy 1583 PhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGln 1602  
 Db 6726 TTCAGCATCTCTCCAGAGCAAAAGCGACTCCACATAGAGACTACTCTGTTTCTCAG 6785  
 Qy 1603 AlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAspAsn--- 1621  
 Db 6786 ACAACACTTGACCAAGTATTGTGAACCTTGGCAAGGACCAAAAGTGAATGATGACCACTTA 6845  
 Qy 1622 -----SerCysGlyThrLeuAsnSerThr 1629  
 Db 6846 AAGACCTCTCATTAACAAAACACAGACAGTAGTGGAGTTGCAAGTTCTCACATCTTTT 6905

Qy 1630 LeuTyrTyrGluArgThrGlnGluAsp-----ArgValValPhe 1642  
 Db 6906 CTACAGGATGAGAAAGTGAAGAAAGCTATGATGAAGAATCCTGTTC 6953  
 RESULT 14  
 US-10-796-307-77  
 ; Sequence 77, Application US/10796307  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CLO01509  
 ; CURRENT APPLICATION NUMBER: US/10796,307  
 ; CURRENT FILING DATE: 2004-03-10  
 ; NUMBER OF SEQ ID NOS: 4201  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 77  
 ; LENGTH: 10405  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-796-307-77  
 Alignment Scores:  
 Pred. No.: 3,678-130 Length: 10405  
 Score: 1439.00 Matches: 485  
 Percent Similarity: 41.88% Conservative: 309  
 Best Local Similarity: 25.58% Mismatches: 632  
 Query Match: 17.08% Indels: 470  
 DB: Gaps: 64  
 US-10-090-458-5 (1-1642) x US-10-796-307-77 (1-10405)  
 Qy 89 SerSerIleMetGlnLysValSerThrAspHisLeuProaspValIleIleThrGluGlu 108  
 Db 1808 TCTCGTTCATCGAGTGTGTCAACCTGAACAAAGTCT---GAACCCATAGCAACAGATC 1864  
 Qy 109 TyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLysProSerAsnPheValGly 128  
 Db 1865 TGGCTCATCAACAGTCCATG-----GAGCTGTGTGATGAGAGAAAGTCTGGCTGGT 1918  
 Qy 129 ValValPhe-----LysAspSerMetSerTyrGlu 138  
 Db 1919 ATTGTGTTCACTGGAAATTAATCCAGGAGATTCAGCTGCCCATCATGTCAAGTACAAG 1978  
 Qy 139 LeuArgPhePheProaspMetIle-----ProValSerSerIleTyrMetAsp 154  
 Db 1979 ATCCGAATGGACATTCAGCAATGTGAGAGGACAAATAAATCAAGAGTGGTACTGGAC 2038  
 Qy 155 -----SerArgAlaGlyCysSerLysSerCysGluAlaAlaGlnTyrTrpSerSerGly 172  
 Db 2039 CCTGCTCTCTCAGCT-----GACCCCTTTGAGGACATCGGTACGCTCTSGGGGGC 2089  
 Qy 173 PheThrValLeuGlnAlaSerIleAspAlaIleIleGlnLeuLysThrAsnValSer 192  
 Db 2090 TTGCGCTACTTTCGCGAGTGTGTGAGCAGCAATCATCAGGGTGTGCTGACGGGACCC--- 2146  
 Qy 193 LeuTyrLysGluLeuGluSerThrLysAlaValIleMetGlyGluThrAlaValAlaGlu 212  
 Db 2147 -----GAGAAGAAAACCTGGTCTCTATATGCACACAGATG----- 2179  
 Qy 213 IleAspThrPheProArgGlyValIleLeuIleTyrLeuValIleAlaPheSerProPhe 232  
 Db 2180 -----CCCTATCCCTGTATGATGACATCTTCTGCGGGTGTGATGAGCCGCTCAATG 2233  
 Qy 233 GlyTyrPheLeuAlaIle-----HisIleVal 241  
 Db 2234 CCCCTTTCATGACGCTGGCTGGATTTACTCAGTGGCTGTGATCATCAAGGGCATTTGTG 2293  
 Qy 242 AlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThrAla 261  
 Db 2294 TATGAGAGGAGGCGCGCTGAAAGAGACCATGGGATCATGGGCTCTGGACACAGCATM 2353



Db 1761 ATTGTGTTCACTGGAATTACTCCAGGAGCAGVATTGAGCTGCCCATCATGTCAAGTACAAG 1820  
Qy 139 LeuArgPhePheProaspMetIle-----ProValSerSerIleThrMetAsp 154  
Db 1821 ATCCGAATGACATTTGACAAATGTGGAGAGACAAATAAATCAAGATGGTACTGGGAC 1880  
Qy 155 -----SerArgAlaGlyCysSerLysSerCysGluAlaAlaGlnTyrTrpSerSerGly 172  
Db 1881 CCTGGTCTCTGAGCT-----GACCCCTTTGAGGACATGGGTACGTCTSGGGGGC 1931  
Qy 173 PheThrValLeuGlnAlaSerIleAspAlaAlaIleIleLeuLeuLeuLeuLeuValSer 192  
Db 1932 TTGCGCTACTTTCGRGATGTGGAGAGCAGCAATCATCAGGGTGTCTGACGGGCACC----- 1988  
Qy 193 LeuTrpLysGluLeuGlnSerThrLysAlaValIleMetGlyLeuThrAlaValGlu 212  
Db 1989 -----GAGAAAGAAACTGGTGTCTATATGCAACAGATG----- 2021  
Qy 213 IleAspThrPheProArgGlyValIleLeuLeuLeuValIleAlaPheSerProPhe 232  
Db 2022 -----CCCTATCTCTGTTASGTTGATGACATCTTCTGCGGTGATGAGCCGGTCAATG 2075  
Qy 233 GlyTyrPheLeuAlaIle-----HisIleVal 241  
Db 2076 CCCTCTTCATGAGCTGGCTGGATTTACTAGTGGCTGTGATCATCAAGGCATYGTG 2135  
Qy 242 AlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThrAla 261  
Db 2136 TATGAGAAGAGGCGCGCTGAAAGAGACCATCGGATCATCGGCTGATGGCTGCAACACATM 2195  
Qy 262 PheTrpLeuSerTrpValLeuLeuTyrThrSerIlePheLeuMet-----Ser 278  
Db 2196 CTCTGGTTTAGCTGTTCAAT-----AGTAGCTCATCTCTCTCTGTGAGCGTGGC 2249  
Qy 279 LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProLysSerSerIleVal 298  
Db 2250 CTGTAGTGTCTCATCTCCGAAATAGGAACCTCTG---CCCTACAGTGCATCCACGCTG 2306  
Qy 299 IlePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThr 318  
Db 2307 GTGTTTGTCTCTGTCCTGTCGGTGTGTGTCGTGGTGACAACTCTGTCAGTCTCTCGATTAGC 2366  
Qy 319 ProLeuPheLysLysSerLysHisVal-----GlyIleValGluPhePheVal 334  
Db 2367 ACATCTCTCCAGAGCAACCTGGCAGACCTGTGGGGGCATCATCTACTTTCACGCTG 2426  
Qy 335 ThrValAlaPhe-----GlyPheIleGlyLeuMetIleLeuLeu 347  
Db 2427 TACCTGCCCTAGCTCTGTGTGTGGCATGGCAGGACTACTGGGCTTCACATCAAGATC 2486  
Qy 348 IleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCysThrPhe 367  
Db 2487 TTCGCTAGC-----CTGCTGTCTCCT-----GTGGCTTTT 2516  
Qy 368 ValIleGlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSerPheSer 387  
Db 2517 GGGTTTGGCTGTGAGTACTTTCCTTTTGGAGAGCAGGGCATTTGGATGCGAGTGGGAC 2576  
Qy 388 AsnLeuThrAlaGlyPro-----TyrProLeuIleIleThrIleLeuMet 402  
Db 2577 AACCTGTTTGAGAGTCTCTGTGGAGGAAGATGGCTTCAATCTCACCATTTCGRCTCCATG 2636  
Qy 403 LeuThrLeuAsnSerIlePheThrValLeuLeuAlaValTyrLeuAspGlnValIlePro 422  
Db 2637 ATGCTGTTTACACCTTCTCTATGGGGTGATGACCTGGTACATTGAGGCTGCTTTTCCA 2696  
Qy 423 GlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp----- 440  
Db 2697 GGCACGTACGGAATTCACAGCCCTGTGTATTTTCTTGACACCAAGTCTTACTGTTTGGC 2756  
Qy 441 -----SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsn 454  
Db 2757 RAGGAAAGTGATGAGAAGAGCCACCTCTGGTTTCCAAACAGAAAGAAATTCAGAA----- 2810

Qy 455 ValAsnGlyAsnIleSerPheSerGluIleIleLeuProValSerSerGluPheValGly 474  
Db 2811 -----AATCTGCATGGAGGAG-----GAACCCACCCACTTG----- 2840  
Qy 475 LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnVal 494  
Db 2841 AAGCTGGCGGTGTCATTCAGAACCTGTTAAAGTCTAC-----YGAGATGGGATGAAGGTG 2897  
Qy 495 GluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGly 514  
Db 2898 ---CCTGTCGATGCCCTGGCACTGAATTTTATGAGGCCAGATCACTCTCTTCTGGGC 2954  
Qy 515 HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysProSer 534  
Db 2955 CACAATGAGYGGGAAAGAGACACCATGTCAATCTGACCCGGTGTTCCTCCCGCCACC 3014  
Qy 535 AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
Db 3015 TCGGGCACYGCTACATCTCTGGGAAAGACATTCGC-----TCTGAGATGAGCACCATC 3068  
Qy 555 ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
Db 3069 CGGCAGAACCTGGGGGTCTGTCCCCAGCATAAAGCTGTGTTGACATGCTGACTGTGCAA 3128  
Qy 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGlu 594  
Db 3129 GACACATCTGGTCTATGCGCGCTTGAAAGGCTCTCTRAGAGCACGTGAAGGCGGAG 3188  
Qy 595 ValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLys---AspAsnGlnAlaLys 613  
Db 3189 ATGAGCAGATGGCCCTGGATGTTGGTTTGGCATCAAGCAAGCTGAAAGCAAAACAGC 3248  
Qy 614 LysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnPro 633  
Db 3249 CAGCTGTGAGTGAATGCGAGAAAGCTATCTGTGCCCTTGGCTTTGCGGGGATCT 3308  
Qy 634 LysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleVal 653  
Db 3309 AAGTGTGTCTCTGATGAACCCACAGCTGTGTGGACCTTACTCCCGCAGGGAATA 3368  
Qy 654 TrpAsn---LeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPhe 672  
Db 3369 TGGAGTGTCTGCTGAAATACCGACAA---GGCGCGCACCATTAATCTCTCTACACACCAC 3425  
Qy 673 MetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLys 692  
Db 3426 ATGATGAGCGGACGCTCTGGGKACAGGATTGCCATCATCTCCCATGGGAAGCTGTC 3485  
Qy 693 CysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrArgLeuSerMet 712  
Db 3486 TGTGTGGCTCTCTCCTGTTTCTGAAGAACCAAGYGGGAAACAGGCTACTACCTGACCTTG 3545  
Qy 712 ----- 712  
Db 3546 GTCAGAAGATGTGGAATCCTCCTCAGTTCCTGTCAGAAACAGTAGTAGCAGTGTGCA 3605  
Qy 713 TyrIleAspLysTyrCysAla----- 719  
Db 3606 TACCTGAAAAGGAGGACAGTGTTCCTCAGACAGAGTTCCTGATGCTGGCCTGGGAGGAC 3665  
Qy 720 -----ThrGluSerLeuSerSerLeuValLysGlnHis 730  
Db 3666 CATGASAGTGACCGCTGACCATCGATGCTCTGTCTATCTCCACCTCATCAGGAAGCAT 3725  
Qy 731 IleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPhe 750  
Db 3726 GTGCTGAAGCCCGCTGTTGTAAGACATAGGGCATGAGCTGACCTATGTGTGCTGCATAT 3785  
Qy 751 LysAspMetAspLys-----PheSerGlyLeuPheSerAlaLeuAspSerHis---Ser 767  
Db 3786 GARGCTGTAGGAGGGAGCGCTTTTGTGAACTCTTTTCATGAGATTGATGACCGCTCTCA 3845

Qy 1160 Tyr-----Thr 1161  
 Db 5269 ATTAATGGCAGCGTGGCCACTTTGTGCTGGAGCTGTCACYGACAATAAGCTGAATAAT 5328  
 Qy 1162 ILeAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProIleLeuGly 1181  
 Db 5329 ATCMATGATATCTGAAGTCCGCTGTTCTTGTATCTCCACATTTTGTCTGGAGCA 5385  
 Qy 1182 CysLeuIleSerPheIleLys----- 1188  
 Db 5386 GGGCTCATGCATGCTGGAACAAACAGGCAATGGCTGATGCCCTGGAAAGGTTTGGGAG 5445  
 Qy 1189 -----IleSerTyrLysAsnValArgLysAsnValAspThrTyrAsn 1202  
 Db 5446 AATCGCTTGTGCACCAATATCTTGGACTTGGTGGAGAACTCTTCGCC----- 5499  
 Qy 1203 ProTyrAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTyrIle 1222  
 Db 5500 -----ATGGCCGTGGAAGGGTGTCTTCTCTCATTAAGTCTT----- 5538  
 Qy 1223 PheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAspProphe 1242  
 Db 5539 -----CTGATCCAGTAC-----AGATTC 5556  
 Qy 1243 PheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGlu 1262  
 Db 5557 TTC-----ATCAGGCCAGACCTGTAAATGCAAGCTA-----TCTCTCTGAT--- 5601  
 Qy 1263 AspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGln 1282  
 Db 5602 GATGAGATGAAGATGTGAGCGGGAAGACAGAGAATTTCTTGTGTTGGAGCC----- 5655  
 Qy 1283 CysCysGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLys 1302  
 Db 5656 -----CAGATGACATCTTAGAATCAAGGAGTTGACGAGATATAY----- 5697  
 Qy 1303 LysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCys 1322  
 Db 5698 -----AGAAGGAAGCGGAAGCTGCTGTTGACAGGATTTGCGTGGCC 5739  
 Qy 1323 ValLysLysGlyLysGluLeuLeuGlyProAsnGluValArgLysSerThrIle 1342  
 Db 5740 ATTCCTCTGCTGAGTCTTGGCTCTGGGAGTTAATGGGCTGGAATATCACTCACT 5799  
 Qy 1343 IleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyr 1362  
 Db 5800 TTCAGATGTTACAGAGATACCACTGTTACAGAGGAGATGCTTTTCCTT---AACAA 5856  
 Qy 1363 SerSerGluThrSerGluAspAspAspSerLeuLysCysMetGlyTyrCysProGlnIle 1382  
 Db 5857 AATAGTATCTTATCAACATCCATGAAGTACATCAGAACATGGCTGCTGCTCAGTTT 5916  
 Qy 1383 AsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLys 1402  
 Db 5917 GATGCCATCAGAGCTGTTGACTGGGAGAGAACACGTCGAGTCTTTCCTCTTTGAGA 5976  
 Qy 1403 GlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeu 1422  
 Db 5977 GGAGTCCAGAGAAAGAGTGGCAAGGTGTGTGAGTGGCGGATTCGGAAATCGGCGCTC 6036  
 Qy 1423 LysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCys 1442  
 Db 6037 GTCAAGTATGGAGAAATATGCTGTGTAATAGTGGAGGCAACAAACGCAAGCTCTCT 6096  
 Qy 1443 PheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGly 1462  
 Db 6097 ACAGCATGCTTGTATCGGCGGCTCTCTGTGGTGTCTTGTGATGAACCCACCAAGGC 6156  
 Qy 1463 MetAspProLysAlaLysHisMetTyrArgAlaIleArgThrAlaPheLysAsnArg 1482  
 Db 6157 ATGATCCCAACACCGCGGGTCTTGTGGAAATGTGCCCTAAGTGTCTCAAG---GAG 6213  
 Qy 1483 LysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaValCysAspArg 1502

Db 6214 GGGAGATCAGTAGTGTCTTACATCTCATAGTATGAAGAAATGTGAAGCTCTTTCACATAGG 6273  
 Qy 1503 VallalIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSer 1522  
 Db 6274 ATGCATATGTCATGCAAGAGGTTCAGTGTGCTGGAGTTCAGCATCTTAAATAAT 6333  
 Qy 1523 LysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTyrIleGluAsnLeuGlu 1542  
 Db 6334 AGGTTGGAGATGTTTATACATAGTTGTAGTAATAGCAGGG-----TCCAACCCYRGAC 6387  
 Qy 1543 ValAspArgLeuGlnAspGluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSer 1562  
 Db 6388 CTGAAGCTGTCCAGGATTTCTTGGACTTTCATTTCTGGAAGTGTCTTAAAGAGAAA 6447  
 Qy 1563 PheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSer 1582  
 Db 6448 CACCGGAACATGTCATAATACCAAGCTTCCA---TCTTCATTATCTTCTCTGGCCAGGATA 6504  
 Qy 1583 PhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGln 1602  
 Db 6505 TTCAGATCTCTCTCCAGAGCAAAAGCGACTCCACATAGAGACTACTCTGTTCCTCAG 6564  
 Qy 1603 AlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAspAsn--- 1621  
 Db 6565 ACAACACTTGACCAAGTATTGTGAACCTTGGCAAGGACCAAGTGTATGATGATGATGAT 6624  
 Qy 1622 -----SerCysGlyThrLeuAsnSerThr 1629  
 Db 6625 AAAGACTCTTCATTACACAAAACACAGACAGTAGTGAGMTGTGCGATCTTCACATCTTT 6684  
 Qy 1630 LeuTyrTrpGluArgThrGlnGluAsp-----ArgValValPhe 1642  
 Db 6685 CTACAGGATGAGAAAGTGAAGAGCTATGTATGAAGATCTGTGTTT 6732

RESULT 13  
 ; US-10-796-307-78  
 ; Sequence 78, Application US/10796307  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C0001509  
 ; CURRENT APPLICATION NUMBER: US/10796,307  
 ; CURRENT FILING DATE: 2004-03-10  
 ; NUMBER OF SEQ ID NOS: 44201  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 78  
 ; LENGTH: 10247  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-796-307-78

Alignment Scores:  
 Pred. No.: 3,58e-130 Length: 10247  
 Scores: 1439.00 Matches: 485  
 Percent Similarity: 41.88% Conservative: 309  
 Best Local Similarity: 25.58% Mismatches: 632  
 Query Match: 17.08% Indels: 470  
 DB: 6 Gaps: 64

US-10-090-458-5 (1-1642) x US-10-796-307-78 (1-10247)

Qy 89 SerSerIleMetGlnLysValSerThrAspHisLeuProAspValIleIleThrGluGlu 108  
 Db 1650 TCTCGCTTCATGAGTGTGTCACCAAGCTT---GAACCCATGACCAAGAGTC 1706  
 Qy 109 TyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLysProSerAsnPheValGly 128  
 Db 1707 TGGCTCATCAACAGTCCATG-----GAGCTGCTGATGAGGAGGAGTCTGGGCTGGT 1760  
 Qy 129 ValValPhe-----LysAspSerMetSerTyrGlu 138

3148	TGGAGAGCTGCTGCTGAATAACCGACAA---GGCGCGACCAATTATTCTCTCTACACACCAC	3204
Qy	MetAspGluAlaAspIleLeuAlaAspArguYalaValIleSerGlnGlyMetLeuLys	692
3205	ATGGATGAAGCGGACGCTCTGGGKACAGGATTGCCATCATCTCCCATGGGAAGCTGTGC	3264
Qy	CysValGlySerSerMetPheLeuLysSerIleTrpGlyIleGlyTyrArgLeuSerMet	712
3265	TGTGTGGGCTCTCTCCCTGTTTCTGAAGAACCAGYTGGAACAGGCTACTACCTGACCTTG	3324
Qy	-----	712
3325	GTCAAGAAAGATGTGGAATCCTCCTCAGTTCTCTGCAGAAACAGTAGTAGCACTGTGTC	3384
Qy	TyrIleAspLysTyrCysAla-----	719
3385	TACCTGAAAGAGGACAGCTGTTCTCAGACGAGTTCTGATGCTGCCTGGGCGGCAC	3444
Qy	-----ThrGluSerLeuSerSerLeuValLysGlnHis	730
3445	CATGASAGTGACACGCTGACCATCGATGCTCTGCTATCTCCAACTTCATCAGGAAGCAT	3504
Qy	IleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPhe	750
3505	GTGTCACAGCCCGCTGGTGGAAACATAGGCGATGAGCTGACCTATGCTGTGCCATAT	3564
Qy	LysAspMetAspLys-----PheSerGlyLeuPheSerAlaLeuAspSerHis---Ser	767
3565	GARGCTCTAAGGAGGAGCGCTTGTGGAACTCTTTTCATGAGATTGATGACCGCTCTCA	3624
Qy	AsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLys	787
3625	GACCTGGGCATTTCTAGTTATGGCATCTCAGACAGCACCTTGGAGAATAATTCCTCAAG	3684
Qy	LeuGluValGluAlaGluIleAsp-----	795
3685	GTGGCCGAGAGAGTGGGGTGGATCTGAGACCTCAGATGGTACTCTGCCAGCAAGACA	3744
Qy	-----GlnAlaAspTyrSerValPheThrGln-----	804
3745	AACAGCGGGCTTCGGGGCAAGCAGAGCTGCTCTGCCCGTTCTCACTGAAGATGATGCT	3804
Qy	---GlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu	823
3805	GCTGATCCAATGATTTCTGACATAGACCCAGATCC-----AGAGAGACAGACTTG	3855
Qy	LeuIleLeuSerGluThrLysAlaSerLeu---ValSerThrMetSerLeuTyrLysGln	842
3856	CTCAGTGGGATGGATGGCAAAAGGRTCTTACCAGGTGAAGGCTGGAACCTTACACAGCA	3915
Qy	GlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLysSerVal	862
3916	CAGTTTGTGGCCCTTTGTGGAAGAGACTGCTAATTCGACAGCGGATCGGAAGGATTT	3975
Qy	ArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPheLeuVal	882
3976	TTTGCTCAGATTGTTGCCAGCTGTGTTTGTCTGCATTGCCCTGTGTTTCAGCCTGATC	4035
Qy	HisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeu-----	899
4036	-----GTGCCACCTTTGGCAAGTAGTACCACGCTTGGAACTTCAG	4074
Qy	-----TyrPheLeuLysProGlyAspLysProHisLysTyrLys	912
4075	CCCTGGATGTACACGAAACAGTACACATTGTTGAGCAATGATGCTCTGTGAGCAGCGGA	4134
Qy	ThrSerLeuLeuLeu-----	917
4135	ACCCTGGAATCTTTAAAGCCCTTCACCAAGAACCCTGGCTTGGGACCCCGCTGTATGGAA	4194
Qy	-----	917
4195	GGAAACCAATCCACAGACACRCCCTGCGACGCGGGAGGAGAGTAGTGCACTGTCGCCCA	4254

QY	917	-----	917
DB	4255	GTCCCCAGACCATCATGGACCTCTTCCAGATGGGAACCTGGAACTGCAACATGCAAGAACCCCTTCA	4314
QY	918	-----GlnAenSerAlaAsp-	922
DB	4315	CCTGCATGCCAGTGTATACGACGACAAAATCAAGAAGATGCTGCTGTGTGTCTCCCCCAGGK	4374
QY	923	-----SerAspIleSerAspLeuIleSerPhePhe	932
DB	4375	GCAGGGGGGCTCCTCTCCACAAAGAAAACAAACACACTGCAGATATCCTTTCAGGACCTG	4434
QY	933	ThrSerGlnAenIleMetValThrMet-	941
DB	4435	ACAGGAAGAAACATTTTGGATTATCTGGTGAAGACGTATGTGCAGATCATATAGCAGAAAAGC	4494
QY	941	-----	941
DB	4495	TTAAGAACAAAGATCTGGGTGAATGATGTAGGTATGGCGGCTTTTCCCTGGGTGTGCAGT	4554
QY	942	-----IleAsnAspSer-	946
DB	4555	AATACTCAAGCAGCTCTCCGAGTCGAAGAAGTTAATGATGCCAYCAAAACAAATGAAGAAA	4614
QY	947	TyrValSerValAlaProHisSerAlaAlaLeuAsnValMethHisSerGluLysAspTyr	966
DB	4615	CACCTAAGCTGGCCAAAGACAGTTCGCAGATCGATTCTCTCAACAGCTTGGGAAGATTT	4674
QY	967	Val-----PheAlaAlaValPheAsnSerThrMetValTyr	978
DB	4675	ATGACAGCACTGGACACCAAAATAATGTCAAGTGTGTGTTCAATAACAAAGGCTGGCAT	4734
QY	979	SerLeuProIleLeuValAsnIleIle-	987
DB	4735	GCAATCAGCTCTTCCCTGAATGTCATCAACRATGCCATTCCTCGGGCCAACTCGCAAAAG	4794
QY	988	-----SerAsnTyr-----TyrLeuTyrHisLeuAsnValThrGlu	999
DB	4795	GGAGAGAACCCCTAGCCATTATGGAAATTACTGCTTTCAATCATCCCTGAAATCTCAACAAAG	4854
QY	1000	ThrIleGlnIleTyrSerThrProPheGlnGluIleThrAspIleValPheLysIle	1019
DB	4855	CAG---CAGCTCTCAGAGGTGGCTCYGATGACCACATCAGTGGATGTCCTTGTGTCCATC	4911
QY	1020	GluLeuTyrPheGlnAlaAlaLeuLeuGlyIleValThrAlaMetProTyrPhe	1039
DB	4912	TGTGTCATCTTT-----GCAATGTCCTTGGTCCCAACGACGCTTT	4950
QY	1040	AlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGly	1059
DB	4951	GTCCTATTCTTGATCCAGAGCGGGTCAGCAAGCAAAACACCTCGAGTTCATCAGTGGGA	5010
QY	1060	LeuLeuProSerAlaTyrTrpIleGlyGlnAlaValAspIleProLeuPhePheIle	1079
DB	5011	GTGAAGCCTGTCACTACTGCTCTCTAAATTTGTCTGGATATGTGCAATTACGTTGTC	5070
QY	1080	IleLeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThr	1099
DB	5071	CCTGCCACATGGTCATTATCATCTTCATCTGCTTCCAGCAGAAGTCCTATGTGCTCTCC	5130
QY	1100	ValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPhe	1119
DB	5131	ACCAATCTGCCTGTGCTAGCGCCCTTCACTTTTGTGTATGGGTGGTCAATCACACCTCTC	5190
QY	1120	ThrTrpIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTyrSerPhe	1139
DB	5191	ATGTACCCAGGCTCCTTTGTGTTCAAG-----	5217
QY	1140	IleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGly	1159
DB	5218	-----ATCCCCACGACAGCCTATGTGGTGTCTCACCAGCGTCAACCTCTTTCATTGGC	5268

; CURRENT APPLICATION NUMBER: US/10/796,307  
 ; CURRENT FILING DATE: 2004-03-10  
 ; NUMBER OF SEQ ID NOS: 44201  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 75  
 ; LENGTH: 10026  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-796-307-75

## Alignment Scores:

Pred. No.: 3,45e-130 Length: 10026  
 Score: 1439.00 Matches: 485  
 Percent Similarity: 41.88% Conservative: 309  
 Best Local Similarity: 25.58% Mismatches: 632  
 Query Match: 17.08% Indels: 470  
 DB: 6 Gaps: 64

US-10-090-458-5 (1-1642) x US-10-796-307-75 (1-10026)

QY 89 SerSerIleMetGlnLysValSerThrAspHisLeuProAspValIleIleThrGluGlu 108  
 DB 1429 TCTCGTCTCATGGAGTGTGCAACCTGAACAAGCTR---GAACCCATAGCAACAGAGTC 1485  
 QY 109 TyrThrAsnGlnLysGluMetLeuThrSerSerLysSerLysProSerAsnPheValGly 128  
 DB 1486 TGGCTCATCAACAAGTCCATG---GAGCTGCTGGATGAGAGGAAGTTCTGGGCTGGT 1539  
 QY 129 ValValPhe-----LysAspSerMetSerTyrGlu 138  
 DB 1540 ATTGTGTTCTAGGAATTATCCAGGAGATTTGAGCTGCCCATCATGATCAAGTACAAG 1599  
 QY 139 LeuArgPhePheProAspMetIle-----ProValSerSerIleTyrMetAsp 154  
 DB 1600 ATCCGAATGACATTTGACATGTGGAGAGGACAAATAAATCAAGGATGGGTACTGGGAC 1659  
 QY 155 -----SerArgAlaGlyCysSerLysSerCysGluAlaIleIleTyrTrpSerSerGly 172  
 DB 1660 CTGGTCTCGAGCT-----GACCCCTTTGAGGACATGCGGTACGTCTGGGGGGC 1710  
 QY 173 PheThrValLeuGlnAlaSerIleAspAlaAlaIleIleGlnLeuLysThrAsnValSer 192  
 DB 1711 TTGCGCTACTTGCRCGATGTGGTGGAGCGCAATCATCAGGGTGTGACGGGCACC--- 1767  
 QY 193 LeuTrpLysGluLeuGluSerThrLysAlaValIleMetClyGluThrAlaValGlu 212  
 DB 1768 -----GAGAAGAAACTGGTGTCTATATGCAACAGATG----- 1800  
 QY 213 IleAspThrPheProArgGlyValIleLeuIleTyrLeuValIleAlaPheSerProPhe 232  
 DB 1801 -----CCCTATCCCTGTTASGTTGATGACATCTTTCTGGGGTGTGATGCCCGTCAATG 1854  
 QY 233 GlyTyrPheLeuAlaIle-----HisIleVal 241  
 DB 1855 CCCTCTTCATGACGCTGGCTGGATTTACTAGTGGCTGTGATCATCAAGGCGCATYGTG 1914  
 QY 242 AlaGlnLysGlnLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThrAla 261  
 DB 1915 TATGAGAGAGGCGCGCTGAAGAGACCATCGGATATGGCGCTGACACACACATM 1974  
 QY 262 PheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMet-----Ser 278  
 DB 1975 CTCTGGTTTAGCTGGTTCAAT-----AGTAGCTCATCTCTCTTCTGTGAGCGGTGC 2028  
 QY 279 LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleVal 298  
 DB 2029 CTGCTAGTGTGATCCTCGAAGTAGTAGGAACCTCTG---CCCTACAGTATCCACGCGTG 2085  
 QY 299 IlePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThr 318  
 DB 2086 GTGTTTGTCTTCTGTCGCGTTTGTGTGGTGACATCTCTGCAAGTCTCTCTGATTAGC 2145  
 QY 319 ProLeuPheLysSerLysHisVal-----GlyIleValGluPhePheVal 334

DB 2146 ACATCTTCTCCAGAGCCAACTGGCAGACCCCTGTGGGGCATCACTACTTCTCACGCTG 2205  
 QY 335 ThrValAlaPhe-----GlyPheIleGlyLeuMetIleIleLeu 347  
 DB 2206 TACCTGCCCTACGTCCTGTGTGGCATGGCAGACACTACRTGGGCTTTCMCATCAAGATC 2265  
 QY 348 IleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCysThrPhe 367  
 DB 2266 TTCGCTAGC-----CTGCTGTCTCTCT-----GTGGCTTTT 2295  
 QY 368 ValIleGlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSerPheSer 387  
 DB 2296 GGGTTTGGCTGTGAGTACTTTGCCCTTTTTCAGGAGCAGGCGCATTTGAGTGCAGTGGAC 2355  
 QY 388 AsnLeuThrAlaGlyPro-----TyrProLeuIleIleThrIleIleMet 402  
 DB 2356 AACCTGTTTGAGATCCTGTGGAGAGATGGCTTCAATCTCACCACTCTCCTCTCTCCATG 2415  
 QY 403 LeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnValIlePro 422  
 DB 2416 ATGCTGTTTACACCTTCTCTATGGGTGATGACCTGGTACATTTAGGCTGTCTTTCCA 2475  
 QY 423 GlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp----- 440  
 DB 2476 GGCAGTAGTACGAATTCACAGCCCTGTGATTTTCTTCCACCAAGTCTACTGTTTGGC 2535  
 QY 441 -----SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsn 454  
 DB 2536 RAGGAAAGTGTAGAGAGAGCCCTGTTCCAAACAGAGAGAAATTCAGAA----- 2589  
 QY 455 ValAsnGlyAsnIleSerPheSerGluIleIleGluProValSerSerGluPheValGly 474  
 DB 2590 -----ATCTGCATGGAGAG-----GAACCCACCCACTTG----- 2619  
 QY 475 LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnVal 494  
 DB 2620 AAGTCGGCGTGTCCATTCAGAACCTGGTAAAGTCTAC---YGAGATGGGATGAAGGTG 2676  
 QY 495 GluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuGly 514  
 DB 2677 --GCTGCGATGCTGGCACTGAATTTTATGAGGCGCAGATCACTCTCTCTCTGGC 2733  
 QY 515 HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysProProSer 534  
 DB 2734 CACATGAGYGGGAGAGAGACCCATGTCATCTCTACCGGGTGTGTCCTCCCGCCACC 2793  
 QY 535 AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
 DB 2794 TGGGCACYGCCTACATCTCTGGGAAAGACATTCGC-----TCTGAGATGAGCACCATC 2847  
 QY 555 ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
 DB 2848 CGGCAGAACCTGGGGGTCTGTCCCAGCATMAACGTGTGTGTGACATGCTGACTGTGCA 2907  
 QY 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGlu 594  
 DB 2908 GAACACATCTGGTCTATGCCGCTTGAAGGGCTCTCTRAGAGACAGCTGAAGCGCGAG 2967  
 QY 595 ValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLys-----AspAsnGlnAlaLys 613  
 DB 2968 ATGGAGCAGATGGCCCTGGATTTGGTTGGCATCAAGCAAGCTGAAAGCAAAACAAGC 3027  
 QY 614 LysLeuSerGlyGlyGlnLysArgLysSerLeuGlyIleAlaValLeuGlyAsnPro 633  
 DB 3028 CAGTGTGAGTGAAGTGCAGAGAAAGCTATCTGTGGCTTGGCTTGTTCGGGGGATCT 3087  
 QY 634 LysIleLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleVal 653  
 DB 3088 AAGTGTGTCATCTCGGATGAACCCACAGCTGGTGTGACCTTACTCTCCGCGAGGGAATA 3147  
 QY 654 TrpAsn---LeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPhe 672



Db 4876 GGCTCCGATGACACATCAGTGGATGCTCTGTGTCCATCTGTGTCACTTT----- 4927  
Qy 1026 aLeuLeuGlyIlelleValThraMetProPtyrPheAlaMetGluAsnAlaGluAs 1046  
Db 4928 -----GCAATGTCCTTGTCCAGCCAGCTTTGCGATTCCTGATCCAGA 4974  
Qy 1046 nHisLysIleLeAlaTyThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyTr 1066  
Db 4975 GCGGGTCAGCAAGCAACACCTCGAGTTCATCAGTGGAGTGAAGCTGTCTACTGTG 5034  
Qy 1066 pIleGlyGlnAlaValAlaSpIleProLeuPhePheIlelleLeuLeuMetLeuG 1086  
Db 5035 GCTCTCTAAATTTGCTGGATATGCAATTCAGTTGCTCCCTGCCACATCGTCAATAT 5094  
Qy 1086 ySerLeuLeuAlaPheHisTyGlyLeuTyPheTyThrValLysPheLeuAlaVal 1106  
Db 5095 CATCTTCATCTGCTCCAGCAGAAGTCTATGCTCTCCACCAATCTGCTGTAGC 5154  
Qy 1106 lPheCysLeuIleGlyTyValProSerValIleLeuPheTyThrIleAlaSerPheTh 1126  
Db 5155 CCTTCTACTTTGCTGTATGGGTGTCATATCACCTCTCATGTACCCAGCTCTTTGT 5214  
Qy 1126 rPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTySerValAlaLe 1146  
Db 5215 GTTCAAG-----ATCCCCAGCAC 5232  
Qy 1146 uAlaCysIleAlaIleThrGluIleThrPheMetGlyTy-----ThreAlaTh 1164  
Db 5233 AGCCTATGTGTCTCCAGCGGTGAACCTCTTCATTGGCATTAATGGCAGCGTGCCAC 5292  
Qy 1164 x-----IleLeuHisTy 1168  
Db 5293 CTTTGTCTGGAGCTGTTCCAGCAATAGCTGAATAATATCAATGATCTCTGAAGTC 5352  
Qy 1168 rAlaPheCysIlellelleProleTyProLeuLeuGlyCysLeuIleSerPheIleTy 1188  
Db 5353 CGTCTTC---TTGATCTTCCACATTTTTCCTGGGACGAGGCTCATCGCATGTGAA 5409  
Qy 1188 g-----11 1189  
Db 5410 AACCCAGCAATGCTGATGCCCTGGAAAGTTGGGAGAATCGTTTGTGTCACCAT 5469  
Qy 1189 eSerTrpLysAsnValArgLysAsnValAspThrTyAsnProTrpAspArgLeuSerVa 1209  
Db 5470 ATCTTGGAGCTGTGTGGGAGCAACCTCTTCGCC-----ATGGCCGT 5511  
Qy 1209 lAlaValIleSerProTyLeuGlnCysValLeuTrpIlePheLeuGlnTyTrG 1229  
Db 5512 GGAAGGGGTGTGTCTTCTTCATCTGTT-----CTGATCCAGTAC 5554  
Qy 1229 uLysLysTyGlyGlyArgSerIleArgLysAspProphePheArgAsnLeuSerThry 1249  
Db 5555 -----AGATTCTTC-----ATCAGGCCAG 5574  
Qy 1249 sSerLysAsnArgLysLeuProGluProProAspAsnGluAspGluAspVally 1269  
Db 5575 ACCTGTAATGCAAGCTA-----TCTCTCTGAAT---GATGAAGATGAAGATGTGAG 5625  
Qy 1269 sAlaGluArgLeuLysVallyGluLeuMetGlyCysGlnCysCysGluLysProse 1289  
Db 5626 GCGGGAAGACAGAGAAATCTTGATGTGGAGGC-----CAGAATGACAT 5670  
Qy 1289 rIleMetValSerAsnLeuHisLysGluTyTrpAspLysLysAspPheLeuLeuSerAr 1309  
Db 5671 CTTAGAATCAAGGAGTTCAGGAAGATATAT-----AG 5703  
Qy 1309 gLysVallyLysValAlaThryIleSerPheCysVallyLysLysGlyGluIleLe 1329  
Db 5704 AAGGAGCGGAGGCTGCTGTGACAGGATTTGGTGGCATTCCTCTGTGTGATGCTT 5763  
Qy 1329 uGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThreIleAsnIleLeuValGlyAs 1349  
Db 5764 TGGGCTCTGGAGTTATGGGCTGGAAATCATCACTTTCAAGATGTTAAACAGAGA 5823

## RESULT 12

US-10-796-307-75  
; Sequence 75, Application US/10796307  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001509

Qy 1349 pIleGluProThrSerGlyGlnValPheLeuGlyAspTyTrpSerSerGluThrSerGluAs 1369  
Db 5824 TACCACCTGTTACCAGAGGAGATGCTTCTCT---AACAGAAATAGTATCTTATCAACAT 5880  
Qy 1369 pAspAspSerLeuLysCysMetGlyTyCysProGlnIleAsnProLeuTrpProAspTh 1389  
Db 5881 CCATGAAGTACATCAACAGCATGGCTACTCCCTCAGTTGATGATGATGATGATGATGAT 5940  
Qy 1389 rThrLeuGlnGluHisPheGluIleTyGlyAlaValLysGlyMetSerAlaSerAspMe 1409  
Db 5941 GACTGGGAGAGAACACACGTGGAGTCTTTGCTCTTTGAGAGGAGTCCACAGAGAAGAGT 6000  
Qy 1409 tLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysTh 1429  
Db 6001 TGGCAAGGTGGTGGAGTGGCGCATTCGAAACCTGGGCTCTGTAAGTATCGAGAAATA 6060  
Qy 1429 rValLysLysLeuProAlaGlyIleLysArgLysLeuLysCysPheAlaLeuSerMetLeuG 1449  
Db 6061 TGCTGGTAACATATAGTGGAGGCAACCAACGCAAGCTCTCTACAGCCATGGCTTTGATCGG 6120  
Qy 1449 yAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysG 1469  
Db 6121 CGGCGCTCTGCTGTGTCTTCTGATGAACCCACACAGGATGGATCCCAAGCCCGCG 6180  
Qy 1469 nHisTrpArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeuTh 1489  
Db 6181 GTTCTGTGAATTTGCCCTAAGTGTGTCAAG---GAGGGGAGATCAGTAGTGTCTTAC 6237  
Qy 1489 rThrHisTyMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSerG 1509  
Db 6238 ATCTCATAGTATGAAGAATGTGAAGCTCTTTCACCTAGGATGGCAATCATGTGTCATGG 6297  
Qy 1509 yGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyTrp 1529  
Db 6298 AAGTTGAGTGTCTTGGAGTGTCCAGCATCTCAAAATAAGTTTGGAGATGTTATAC 6357  
Qy 1529 eLeuGluIleLysLeuLysAspTrpIleGluAsnLeuGluValAspArgLeuGlnArgG 1549  
Db 6358 AATAGTTGTAGGAATAGCAGGG-----TCCAAACCCGACCTGAAGCTGTCCAGGATTT 6411  
Qy 1549 uIleGlnTyTrpPheProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAlaTy 1569  
Db 6412 CTTTGGATTCATTTCTGGAAGTGTCCMAAGAGAAACACCGGACATGTTACATA 6471  
Qy 1569 rLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluAl 1599  
Db 6472 CCAGCTTCCA---TCTTCATTATCTTCTGGCCAGGATATTCAGCATCTCTCCAGAG 6528  
Qy 1589 aLysHisAlaPheAlaIleGluGluTyTrpSerPheSerGlnAlaThrLeuGluGlnValPh 1609  
Db 6529 CAAAAGCGACTCCACATAGAGACTCTCTGTTCTTCAGACAAACACTTGACCAAGTATT 6588  
Qy 1609 eValGluLeuThrLysGluGlnGluGluAspAsn----- 1621  
Db 6589 TGTGAACCTTGCACAGGACCAAGTATGATGATGATGATGATGATGATGATGATGATGAT 6648  
Qy 1622 -----SerCysGlyThrLeuAsnSerThrLeuTrpGluArgThrG 1636  
Db 6649 AAACACAGACAGTAGTGGAGCTTGCAGTGTCTCACATCTTTCTCAGGATGAGAAGTGAA 6708  
Qy 1636 nGluAsp-----ArgValAlaPhe 1642  
Db 6709 AGAAAGCTATGATCAAGAAATCCTGTTCC 6736



RESULT 11  
US-10-465-498-96  
; Sequence 96 Application US/10465498  
; GENERAL INFORMATION:  
; APPLICANT: DENEFE, Patrice  
; APPLICANT: ROSIER-MONTUS, Marie-Francoise  
; APPLICANT: ARNOULD-REGUIGNE, Isabelle  
; APPLICANT: PRADES, Catherine  
; APPLICANT: NAUDIN, Laurent  
; APPLICANT: LEMOINE, Cedrine  
; APPLICANT: DIVERGER, Nicolas  
; APPLICANT: ASSMANN, Gerd  
; APPLICANT: RUST, Stephan  
; APPLICANT: FUNKE, Harold  
; APPLICANT: BREWER, Bryan  
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS CORRESPONDING TO HUMAN GENE ABC1  
; FILE REFERENCE: ST990205CN71  
; CURRENT APPLICATION NUMBER: US/10/465,498  
; CURRENT FILING DATE: 2003-06-18  
; PRIOR APPLICATION NUMBER: FR99 07684  
; PRIOR FILING DATE: 1999-06-17  
; PRIOR APPLICATION NUMBER: PCT/FR00/01595  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: US10/018,714  
; PRIOR FILING DATE: 2001-12-17  
; PRIOR APPLICATION NUMBER: US10/269,780  
; PRIOR FILING DATE: 2002-10-11  
; NUMBER OF SEQ ID NOS: 159  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 96  
; LENGTH: 9495  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (6763)..(6763)  
; OTHER INFORMATION: n is a, c, g, or t  
US-10-465-498-96  
Alignment Scores:  
Pred. No.: 3,76e-132 Length: 9495  
Score: 1458.50 Matches: 491  
Percent Similarity: 43.33% Conservative: 328  
Best Local Similarity: 25.98% Mismatches: 613  
Query Match: 17.31% Indels: 458  
DB: 6 Gaps: 68  
US-10-090-458-5 (1-1642) x US-10-465-498-96 (1-9495)  
Qy 89 SerSerleMetGlnLysValSerThrAspHisLeuProaspVallelleThrGluGlu 108  
Db 1435 TCTCGCTTCATGGAGTGTGCAACCTGACCAAGCTA---GAACCCATAGCAACAGAGTC 1491  
Qy 109 TyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLysProSerAsnPheValGly 128  
Db 1492 TGGCTCATCAACAGTCCATG-----GAGCTGCTGATGAGAGGAGTCTGGGCTGGT 1545  
Qy 129 ValValPhe-----LysAspSerMetSerTyrGlu 138  
Db 1546 ATTGTGTTCTACTGGAATTACTCCAGGCAGCATTCAGCTGCCCATCATGTCACAGTACAAG 1605  
Qy 139 LeuArgPhePheProaspMetIle-----ProValSerSerIleTyrMetAsp 154  
Db 1606 ATCCGAATGACATTGACATGTGGAGGACAAATAAATCAAGATGGGTACTGGAC 1665  
Qy 155 -----SerArgAlaGlyCysSerLysSerCysGluAlaAlaGlnTyrTrpSerSerGly 172  
Db 1666 CCTGGTCTCGAGCT-----GACCCCTTTGAGGACATCGGTACGTCTCGGGGGGC 1716  
Qy 173 PheThrValLeuGlnAlaSerIleAspAlaAlaIlelleGlnLeu-----Lys 188  
Db 1717 TTCCTACTTGCCAGGATGTGGGACAGGCATCATCAGGGTGTCTCGGGCACCGAGAA 1776

Qy 189 ThrAsnValSerLeuTrpLysGluLeuGluSerThrLysAlaVallelleMetGlyGluThr 208  
Db 1777 GAAAACCTGGTGTCTATAT-----GCAACAGATGCCCTA----- 1809  
Qy 209 AlaValValGlu-leaspThrPheProArgGlyVal-----lleLeuIleTyrLe 225  
Db 1810 TCCCTGTTACGTTGATGACATCTTTCTGGGGTGATAGCGGGTCAATGCCCCCTCTTCAT 1869  
Qy 225 uValleAlaPheSerProPheGlyTyrPheLeuAlaIle-----HisleValAl 242  
Db 1870 GACGCTGCGCTGGATT-----TACTCAGTGGCTGTGATCATCAAGGGCATCGTGTA 1920  
Qy 242 agluLysGluLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 262  
Db 1921 TGAGAGAGGACGCGCTGAAGAGACATCGCGGATCATGGGCTTGACACACATCCT 1980  
Qy 262 eTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMet-----SerLe 279  
Db 1981 CTGGTTTAGCTGGTTCATT-----AGTAGCCTCATCTCTCTTCTGAGCGCTGCGCT 2034  
Qy 279 uLeuMetAlaVallelleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleVal 299  
Db 2035 GCTAGTGGTCATCTCTGAAGTTAGGAACCTGCTG---CCCTACAGTATCCACGCTGGT 2091  
Qy 299 ePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrPr 319  
Db 2092 GTTTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2151  
Qy 319 oLeuPheLysLysSerLysHisVal-----GlylleValGluPhePheValTh 335  
Db 2152 ACTCTTCTCCAGAGCCAACTGGCAGCAGCGCTGGGGGCGATCATCTACTTCACGCTGTA 2211  
Qy 335 rValAlaPhe-----GlyPheIleGlyLeuMetIleLeuIle 348  
Db 2212 CTGCCCCCTACGTCCTGTGTGGCATGGCAGGACATGCTGGGCTTTCACACTCAAGATCTT 2271  
Qy 348 eGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCysThrPheVa 368  
Db 2272 CGCTAGC-----CTGCTCTCTCTCT-----GTGGCTTTTGG 2301  
Qy 368 lIleGlyIleAlaGlnValMetHisLeuGluAaspPheAsnGluGlyAlaSerPheSerAs 388  
Db 2302 GTTGGCTGTGATGACTTTCCTTTTGGAGCAGGCGCATTTGGAGTGGAGTGGGACAA 2361  
Qy 388 nLeuThrAlaGlyPro-----TyrProLeuIlelleThrleleMetle 403  
Db 2362 CCTGTTTGAGAGTCTCTGGAGGAAGATGGCTTCAATCTCACCACCTCGGTCTCCATGAT 2421  
Qy 403 uThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAaspGlnVallelleProG 423  
Db 2422 GCTGTTTGACACCTTCTCTATGGGGTGATGACCTGTGATATGAGGCTGTCTTCCAGG 2481  
Qy 423 yGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp----- 440  
Db 2482 CCAGTACGGAATTCACAGGCCCTGGTATTTTCTTGCACCAAGTCTACTGTTTGGCGA 2541  
Qy 441 -----SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnVa 455  
Db 2542 GGAAGATGATGAGAGAGCCACCTGGTTCACACAGAGAGATATCAGAA----- 2593  
Qy 455 lAsnGlyAsnIleSerPheSerGluIlelleGluProValSerSerGluPheValGlyLy 475  
Db 2594 -----ATCTGCATGGAGGAG-----GAACCCACCCACTTG-----AA 2625  
Qy 475 sGluAlaIleArgIleSerGlylleGlnLysThrTyrArgLysLysGlyGluAsnValG 495  
Db 2626 GCTGGGCTGTCCATTCAGAACCTGGTAAAGTCTACCGA---GATCGGATGAAGGTG-- 2680  
Qy 495 uAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHi 515  
Db 2681 -GCTGTCATGGCTGCGCACTGAATTTTATGAGGGCCAGATCATCTCTCTCTCTCTCTCTCT 2739

Db 4879 CGGCGCAACCTGCAAAAGGAGAGAACCCCTAGCATATGGAATTACTGTTTCAATCAT 4938  
 Qy 994 HisLeuAenValThrGluThrIleGlnIleTrpSerThrProPheGlnGluIleThr 1013  
 Db 4939 CCCCTGAATCTCACAAGCAG---CAGCTCTCAGAGGTGGCTCCGATGACCATCAGTG 4995  
 Qy 1014 AspIleValPheLysIleGluLeuTyPheGlnAlaAlaLeuGlyIleIleValThr 1033  
 Db 4996 GATGCTCTGTGTCATCTGTGTCATCTTT-----GCAATGTC 5034  
 Qy 1034 AlaMetProProTyPheAlaMetGluAenAlaGluAenHisLysIleLysAlaTyThr 1053  
 Db 5035 TTGCTCCCGCAGCTTTGTCGATTCTCTGATCCAGGAGCGGTTCAGCAAGCAAAACAC 5094  
 Qy 1054 GlnLeuLysLeuSerGlyLeuLeuProSerAlaTyTrpIleGlyGlnAlaValAla 1073  
 Db 5095 CTGCAGTTCATCAGTGAGTGAAGCCCTGTCATCTACTGCTCTCTAATTTGCTGGGAT 5154  
 Qy 1074 IleProLeuPhePheIleIleLeuIleMetLeuGlySerLeuLeuAlaPheHisTy 1093  
 Db 5155 ATGTGCAATACGTTGCTCCCGCCACACTGGTCATTATCATCTCTGCTCCAGCAG 5214  
 Qy 1094 GlyLeuTyPheTyThrValLysPheLeuAlaValPheCysLeuIleGlyTyVal 1113  
 Db 5215 AAGTCTCATGTCTCTCCCAATCTGCTGCTGCTAGCCCTTCTACTTTTGTGTATGG 5274  
 Qy 1114 ProSerValIleLeuPheThrTyIleAlaSerPheThrPheLysLysIleLeuAenThr 1133  
 Db 5275 TGGTCAATCACACTCTCATGATCCAGCTCTCTTGTGTTCAG----- 5319  
 Qy 1134 LysGluPheTrpSerPheIleTySerValAlaLeuAlaCysIleAlaIleThrGlu 1153  
 Db 5320 -----ATCCGAGCAGCAGCTATGTGTGTCTCAGCAGC 5352  
 Qy 1154 IleThrPhePheMetGlyTy-----ThrIleAlaThr----- 1164  
 Db 5353 GTGAACCTCTCATGTGCAATATGCGAGCGGTGGCCACCTTTGTGTGGAGCTGTTCACC 5412  
 Qy 1165 -----IleLeuHisTyTrpAlaPheCysIleIleIlePro 1175  
 Db 5413 GACAATAAGCTGAATAATATCAATGATATCTGAAAGCCGTGTC---TTGATCTTCCCA 5469  
 Qy 1176 IleTyProLeuLeuGlyCysLeuIleSerPheIleLys----- 1188  
 Db 5470 CATTTTTCCTGGGACGAGGCTCATCGACATGCTGAAACACGAGCAATGGCTGATGCC 5529  
 Qy 1189 -----IleSerTrpLysAenValArgLys 1196  
 Db 5530 CTGGAAGGTTTGGGGAGATCGCTTTGTCTCACCATTATCTTGGGACTTGGTGACGCA 5589  
 Qy 1197 AsnValAspThrTyPheAsnProTyPheArgLeuSerValAlaValIleSerProTyLeu 1216  
 Db 5590 AACCTCTTCGCC-----ATGCGCGTGAAGGGTGGTGTCTTCTTC 5631  
 Qy 1217 GlnCysValLeuTrpIlePheLeuGlnTyTyGluLysLysTyGlyGlyArgSer 1236  
 Db 5632 ATACTGTT-----CTGATCCAGTAC----- 5652  
 Qy 1237 IleArgLysAspProPhePheArgAenLeuSerThrLysSerLysAenArgLysLeuPro 1256  
 Db 5653 -----AGATTTCTTC-----ATCAGCCCGCAGCCTCTAATCAAGCTA--- 5691  
 Qy 1257 GluProProAspAenGluAspGluAspValLysAlaGluArgLeuLysValLys 1276  
 Db 5692 ---TCTCTCTGAT---GATGAAGATGAAGATGTGAGCGGGAAGAGCAGAAATCTT 5745  
 Qy 1277 GluLeuMetGlyCysGlnCysCysGluGluLysProSerIleMetValSerAenLeuHis 1296  
 Db 5746 GATGGTGGAGGC-----CAGATGACATCTTAGAAATCAAGGAGTTGAGC 5790  
 Qy 1297 LysGluTyPheAspLysLysAspPheLeuLeuSerArgLysValLysValAlaThr 1316

Db 5791 AAGATATAT-----AGAAAGAAAGCGGAGCCCTGCTGTT 5823  
 Qy 1317 LysTyIleSerPheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGly 1336  
 Db 5824 GACAGATTTGCTGGGCATCTCTCTGCTGAGTCTTTGGGCTCTCTGGAGTTAATGG 5893  
 Qy 1337 AlaGlyLysSerThrIleAlaAsnIleLeuValGlyAspIleGluProThrSerGlyGln 1356  
 Db 5894 GCTGGAATAATCATCACTTTCAAGATGTTACAGAGAGATACCATGTTTACAGAGAGAT 5943  
 Qy 1357 ValPheLeuGlyAspTySerSerGluThrSerGluAspAspSerLeuLysCysMet 1376  
 Db 5944 GCTTTTCTT---ACAGAAATAGTATCTTATCAAAACATCCATGAAGTACATCAACAATG 6000  
 Qy 1377 GlyTyCysProGlnIleAsnProLeuTrpProSerThrThrLeuGlnGluHisPheGlu 1396  
 Db 6001 GCTACTGCTCTCAGTTTGTATGCCATCACAGAGCTGTGATGGAGAGAAACATGTTGAG 6060  
 Qy 1397 IleTyGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIle 1416  
 Db 6061 TTTCTTTGCCCTTTTGAGAGGAGTCCAGAGAAAGAGTGGCAAGTTGGTGGTGGCG 6120  
 Qy 1417 ThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGly 1436  
 Db 6121 ATTCGGAACCTGGGCTCTGTGAAGTATGGAGAAATATGCTGTAACTATAGTGGAGGC 6180  
 Qy 1437 IleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAenProGlnIleThrLeuLeu 1456  
 Db 6181 AACAAACCAAGCTCTCTACAGCCATGCTTGTATCGCGGGCTCTCTGTTGTTCTG 6240  
 Qy 1457 AspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArg 1476  
 Db 6241 GATGAACCCACACAGGATGATGCCAAAGCCCGGGGTTCTTGTGGAATGTGCCCTA 6300  
 Qy 1477 ThrAlaPheLysAenArgLysArgAlaAlaIleLeuThrThrHisTyMetGluGluAla 1496  
 Db 6301 AGTGTGTGCAAG---GAGGGGAGATCAGTAGCTGTATCATCTCATAGTATGGAAGATGT 6357  
 Qy 1497 GluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThr 1516  
 Db 6358 GAAGCTCTTTCACCTAGGATGCAATCATGCTCAATGCAAGTTTCAGTGGCTTGGAGAT 6417  
 Qy 1517 ValGlnHisLeuLysSerLysPheGlyLysGlyTyTyPheLeuGluIleLysLeuAsp 1536  
 Db 6418 GTCCAGCATCTAATAAATAGGTTTGGAGATGGTTATACATAGTTGTACGATAGCAGGG 6477  
 Qy 1537 TrpIleGluAenLeuGluValAspArgLeuGlnArgGluIleGlnTyIlePheProAsn 1556  
 Db 6478 -----TCCAAACCCGACCTGAAGCTGTCCAGGATTTCTTTGGACTTGCATTTCTCGGA 6531  
 Qy 1557 AlaSerArgGlnGluSerPheSerSerIleLeuAlaTyLysIleProLysGluAspVal 1576  
 Db 6532 AGTGTTCAAAGAGAAACACCGGAACATGCTCAATACAGCTTCCA---TCTTCATTA 6588  
 Qy 1577 GlnSerLeuSerGlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGlu 1596  
 Db 6589 TCTTCTCTGGCCAGGATATTCCAGCATCTCTCCAGAGCAAAAGCGACTCCACATAGAA 6648  
 Qy 1597 GluTySerPheSerGlnAlaThrIleuGluIleValPheValGluLeuThrLysGluGln 1616  
 Db 6649 GACTACTCTGTTTCTCAGACCAACTTGACCAAGTATTTGTGAACCTTTGGCAAGGACCAA 6708  
 Qy 1617 GluGluGluAspAen-----SetCys 1623  
 Db 6709 AGTGATGATGACACTTAAAGACCTCTCATACAAAACACAGACAGTAGTCGCGTT 6768  
 Qy 1624 GlyThrLeuAenSerThrLeuTrpTrpGluArgThrGlnGluAen-----ArgVal 1640  
 Db 6769 GCAGTCTTCATCTCTTTCTACAGGATGAGAAAGTGAAGAAAGCTATGTATGAAGATC 6828  
 Qy 1641 ValPhe 1642  
 Db 6829 CTGTTTC 6834

Qy	483	IleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPhe	502	Qy	795	-----	795
Db	2746	CTGTAAGATCTACCA---GATGGATGAAGTG---CTCTCGATGCCCTGGCACTG	2799	Db	3811	GCTGAGACCTCAGATGGTACCTTCCAGCAAGACGAAACAGCGCGGCTTCGGGGACAAG	3870
Qy	503	AspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThr	522	Qy	796	GlnAlaAspTyrSerValPheThrGln-----GlnProLeuGluGluGluMet	811
Db	2800	AAATTTTATAGGGCCAGATCACCTCTCTCTGGCCACATATGAGCGGGAGAGACGCC	2859	Db	3871	CAGAGCTCTCTCCCGCTTCAGTGAAGATGATCTGCTGATCCAAATGATTCGACATA	3930
Qy	523	LeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIleTyrGly	542	Qy	812	AspSerLysSerPheAspGluMetGluGlnSerLeuLeuLeuLeuLeuLeuLeuLeu	831
Db	2860	ACCATGTCAATCTGACCGGGTGTTCCTCCCGGACCTCGGCGACCGCTACATCTCTGGGA	2919	Db	3931	GACCCAGAATCC-----AGAGACACAGACTTGCTCAGTGGGATGGCAAGGG	3981
Qy	543	HisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyCysPro	562	Qy	832	SerLeu---ValSerThrMetSerLeuTyrLysGlnMetTyrThrIleAlaLysPhe	850
Db	2920	AAAGACATCGC-----TCTGATGAGCACCATCCGGCAGAACCTGGGGGTCTGTCC	2973	Db	3982	TCTTACCAGGTGAAGGCTGGAACCTTACAGACACAGATTGTGGCCCTTTTGTGAAG	4041
Qy	563	GlnLeuAspIleHisPheAspValIleThrValGluGluAsnLeuSerIleLeuAlaSer	582	Qy	851	HisPhePheThrLeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeu	870
Db	2974	CAGCATAACTGTGTTTGCATGCTGCTGTGTGAGAACACATCTGTTCTATGCCCGC	3033	Db	4042	AGACTGTCTAATTCAGACAGCGAGTCGGAAGGATTTTGTCTCAGATTGTCTTCCAGCT	4101
Qy	583	IleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeu	602	Qy	871	IlePhePheThrValGlnIlePheMetPheLeuVal-----	882
Db	3034	TTGAAGGGCTCTCTGAAGACACGCTGAAGCGGAGATGAGCAGATGGCCCTGGATGT	3093	Db	4102	GTGTTTGTCTGCATTTGCCCTTGTGTTGAGCCTGTGTCGACCTGTGTCGACCTTGGCAAGTACCCC	4161
Qy	603	AspMetGlnThrIleLys---AspAsnGlnAlaLysLysLeuSerGlyGlyGlnLysArg	621	Qy	882	-----	882
Db	3094	GGTTTGGCATCAAGCAAGCTGAAGCAAAACAGCCGCTGTCAGTGGATGCAGAGA	3153	Db	4162	AGCCTGGAACCTCAGCCCTGGATGTACAAACGACAGTACACATTGTGTGACCAATGATGCT	4221
Qy	622	LysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGluPro	641	Qy	882	-----	882
Db	3154	AAGCTATCTGTGGCTTGGCTTTCTCGGGGATCAAGGTTGCTCATTTCTGGATGAACCG	3213	Db	4222	CCTGAGGACACGGGAACCTGGAACTCTTAAACGCTCACCAGAACCCCTGGCTTCGGG	4281
Qy	642	ThrAlaGlyMetAspProCysSerArgHisIleValTyrAsn---LeuLeuLysTyrArg	660	Qy	883	-----	883
Db	3214	ACAGCTGTGTGGACCTTACTCCCGCAGGGGAATATGGAGCTGTCTGTGAATACCGA	3273	Db	4282	ACCCGCTGTATGAAGNAACCAATCCAGACACGCGCTGCCAGCGAGGGAGAGAG	4341
Qy	661	LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla	680	Qy	886	PhelysAsnAlaValValProIleLysLeuValProAspLeuTyr-----	900
Db	3274	CAA---GGCGCACCATTTCTCTACACACACATGATGAAGCGGACGTCCTGGGG	3330	Db	4342	TGGACCACTGCCCGGATTCCTCC---CAGACCATCATGACCTCTTCCAGAATGGGACTGG	4398
Qy	681	AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu	700	Qy	900	-----	900
Db	3331	GACAGGATTCATCTCCATCTCCATGGGAAGCTGTGTGTGGGCTCTCCCTGTTCTG	3390	Db	4399	ACAATGAGAACCTTACCTGCTGATGCGAGTGTAGCAGGACGACAAATCAAGAGATGCTG	4458
Qy	701	LysSerLysTrpGlyIleGlyTyrArgLeuSerMet-----	712	Qy	901	---PheLeuLysProGly-----AspLysProHisLysTyrLysThrSer	914
Db	3391	AAGAACAGCTGGGAACAGGCTACTACCTGACCTTGTGTCAAGAAAGATGTGAATCCTCC	3450	Db	4459	CCTGTGTCTCCCGCAGGGGCGGGGGTCTCTCCCAAGAAACAAACACATGCA	4518
Qy	713	-----TyrIleAspLysTyrCysAla---	719	Qy	915	LeuLeuLeuGlnAsnSerAlaAspSerAspIleSerAsp-----	927
Db	3451	CTCAGTTCTCTCAGAAACAGTAGCAGTGTGTGTATACCTGAAAGAGGAGGACAGTGT	3510	Db	4519	GATATCTCTCAGGACCTGACAGGAAGAAACATTTTCGGATTATCTGGTGAAGACGTATGTG	4578
Qy	719	-----	719	Qy	928	---LeuIleSerPhePheThrSerGlnAsnIleMetValThr-----	940
Db	3511	TCTCAGACAGTCTCTGATGTGGCTGGCGCAGCACCATGAGAGTGCACGCTGACCATC	3570	Db	4579	CAGATCATAGCCAAAGCTTAAAGAACAGATCTGGGTGAATGATTTAGGTATGGCGGC	4638
Qy	720	---ThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGln	738	Qy	941	-----MetIleAsnAspSer	945
Db	3571	GATGCTCTGCTATCTCCAACTCATCAGGAAGCATGTGTCTGAAGCCCGGCTGTGGAA	3630	Db	4639	TTTTCCCTGGGTGTCAGTAATACCTCAAGCACTTCTCCGAGTCAAGAAATTAATGATGCC	4698
Qy	739	GlnAsnAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLys-----Phe	756	Qy	946	-----AspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet	960
Db	3631	GACATAGGCGATGAGTGCATGTCTGCTCCATATGAGCTGCTAAGGAGGAGGCGCTTT	3690	Db	4699	ACCAACCAATGAAGAACACCTTAAGCTGCCAAGGACAGTCTCTGCAGATCGATTCTC	4758
Qy	757	SerGlyLeuPheSerAlaLeuAspSerHis---SerAsnLeuGlyValIleSerTyrGly	775	Qy	961	HisSerGluLysAspTyrVal-----PheAlaAlaValPhe	972
Db	3691	GTGGAATCTCTTATGATGATGATGACCGGCTCTCAGACCTGGGCATTTCTAGTTATGGC	3750	Db	4759	AACAGCTTGGGAAGATTTATGACAGGACTGGACACCAAGAAATAATGTCAAGGTGTGGTTC	4818
Qy	776	ValSerMetThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAsp	795	Qy	973	AsnSerThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeu	992
Db	3751	ATCTCAGACGACCCCTGGGAAGAATAATCTCTCAAGGTGGCCGAGAGAGTGGGGTGGAT	3810	Db	4819	AATAACAGGCGTGGCATGCAATCAGCTCTTCTCTGATGTCTATCAACATGCCATTCTC	4878
				Qy	993	-----Tyr	993



Qy	993	-----TyrHisLeuAsnValThrGluThrIleGlnIleTrpSer	1005
Db	4819	TATGGAATTACTGCTTTCAATCATCCCTCGAATCCACCAAGCAG---CAGCTCTCAGAG	4875
Qy	1006	ThrProPhePheGlnGluIleThrAspIleValPheLysIleGluLeuTyrPheGlnAla	1025
Db	4876	GTGGCTCGGATGACCATCAGTCGATGTCCTGTGTGCCATCTCTGTGATCTTTT---	4929
Qy	1026	AlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAlaMetGluAsnAlaGlu	1045
Db	4930	-----GCAATGTCTCTGCTCCAGCCAGCTTGTGCTATCTCTGATCCAG	4974
Qy	1046	AsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyr	1065
Db	4975	GAGCGGTGACCAAGAAACACCTGCAGTTCATGAGTGAGGAGAGCCTGTCATCTAC	5034
Qy	1066	TrpIleGlyGlnAlaValAspIleProLeuPhePheIleIleLeuIleLeuMetLeu	1085
Db	5035	TGGCTCTCTAATTTTGTCTGGATATGTGCAATTACGTTGTGCCCTGCCACACTGGTCAT	5094
Qy	1086	GlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaVal	1105
Db	5095	ATCATCTCATCTGCTTCCAGCAGAAGTCCTATGTCTCCACCACAAATGCTGCTGCTA	5154
Qy	1106	ValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPhe	1125
Db	5155	GCCCTTCTACTTTTGCTGTATGGGTGTCATACACACTCTCATGTACCCAGCCTCCTTT	5214
Qy	1126	ThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyrSerValAlaAla	1145
Db	5215	GTGTTTCAAG-----ATCCCCAGC	5232
Qy	1146	LeuAlaCysIleAlaIleThrGluIleThrPheMetGlyTyr-----ThrIleAla	1163
Db	5233	ACAGCTATGTGGTGCTCACCGAGTGACCTCTTTCATTGGCATTAATGCGACGGTGGCC	5292
Qy	1164	Thr-----IleLeuHis	1167
Db	5293	ACCTTTGTGTGGAGCTGTTTCACGACAAATAGCTGAATAATATCATGATATCTCTGAAG	5352
Qy	1168	TyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysIleLeuSerPheIle	1187
Db	5353	TCCGTGTTC---TTGATCTTCCACATTTTTCCTGGGACGAGGGCTCATCGACATGGTG	5409
Qy	1188	Lys-----	1188
Db	5410	AAAAACCAAGCAATAGGCTGATCCCTCGAAGAGTTTGGGGAGAATCGTTTGTGTCAACA	5469
Qy	1189	IleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSer	1208
Db	5470	TTATCTTGGGACTTGGTGGGACGAAACCTCTTCGCC-----ATGGGCC	5511
Qy	1209	ValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyr	1228
Db	5512	GTGGAAGGGGGGTGTTCTTCTCTCATATTGTT-----CTGATCCAGTAC---	5556
Qy	1229	GluLysLysTyrGlyGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThr	1248
Db	5557	-----AGATTCTTC-----ATCAGGCC	5574
Qy	1249	LysSerLysAsnArgLysLeuProLysProAspAsnGluAspGluAspVal	1268
Db	5575	AGACCTGTAATGCAAGCTA-----TCTCTCTGAAAT---GATGAAGATGAAGATGTG	5625
Qy	1269	LysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGlyLysPro	1288
Db	5626	AGGCGGGAAGACAGAGAATTTCTTGATGGTGGAGGC-----CAGAATGAC	5670
Qy	1289	SerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSer	1308
Db	5671	ATCTTAGAATCAAGGAGTTGACGAGATATAT-----	5703
Qy	1309	ArgLysValLysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIle	1328

5704	AGAAGGAAGCGGAAGCCTGCTGTGTGACAGATTTCGGTGGCGCATTCTCTCTGTGTAGTGC	5763
1329	LeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThrIleIleAsnIleLeuValGly	1348
5764	TTTGGGCTCTGGAGTTAATGGGGCTGGAAATCATCACTTTCAAGATGTTTACAGGA	5823
1349	AspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGlu	1368
5824	GATACCACTGTTACCAAGAGAGATGCTTTCCTT---AACAGAAATAGTATCTTATCAAA	5880
1369	AspAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTrpProAsp	1388
5881	ATCCATGATGATACATCAGACATGGGCTACTGCGCTCAGTTTCATGCCATCAGAGCTG	5940
1389	ThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAsp	1408
5941	TTGACTGGGAGAGAACACGTCGAGTCTTTTGGCCCTTTTGAGAGAGATCCACAGAA	6000
1409	MetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLys	1428
6001	CTTGGCAGGTTTGGTGTAGTGGCGATTCGGAACTGGGCTCTGTGAAGTATGGAGAAA	6060
1429	ThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeu	1448
6061	TATGCTGGTAATATAGTGGAGGCAACAAACGCAAGCTCTCTACAGCCATGGCTTTG	6120
1449	GlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLys	1468
6121	GGCGGGCTCTGTGGGTTCCTGGATGAACCCACCAAGGATGATGCCAAAGCCCGG	6180
1469	GlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaIleLeu	1488
6181	CGGTCTGTGGAAATTGCGCCCTAAGTGTGTCAAG---GAGGGAGATCAGTAGTGCT	6237
1489	ThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgValaIleMetValSer	1508
6238	ACATCTCATAGTATGGAAGAAATGTGAAGCTCTTTGCATGAGTGAATCATGTGCAAT	6297
1509	GlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyr	1528
6298	GGAAGTTTCAGGTGCTTGGCAGTGTCCAGCATCTAAAAAATAGGTTTGGAGATGTTAT	6357
1529	PheLeuGluIleLysLeuLysAspTrpIleGluAsnLeuGluValAspArgLeuGlnArg	1548
6358	ACAATAGTTGTACGAATAGCAGGG---TCCAAACCCGACCTGAAGCTGTGCCAGAT	6411
1549	GluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSerPheSerIleLeuAla	1568
6412	TTCTTTGGACTTGCAATTTCTTGGAAAGTGTCCAAAGAGAAACACCCGGAACATGCTACAA	6471
1569	TyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluGlu	1588
6472	TACCAGTTTCCA---TCTTCATTATCTTCTGGCCAGGATATTCAGCATCTCTCCCCAG	6528
1589	AlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThrLeuGluGlnVal	1608
6529	AGCAAAAGCGACTCCACATAGAAGACTACTCTGTTTCTCAGACACACATTCGACCA	6588
1609	PheValGluLeuThrLysGluGlnGluGluAspAsn-----	1621
6589	TTTGTGAACCTTGCAGGACCAAGATGATGATGATGATGATGATGATGATGATGATGAT	6648
1622	-----SerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThr	1635
6649	AAAACACAGACAGTAGTGGAGTTGCAGTCTCACATCTTTCTACAGATGAGAAAGTG	6708
1636	GlnGluAsp-----ArgValValPhe	1642
6709	AAAGAAAGCTATGATGAGAAATCCCTGTTCC	6738

RESULT 10  
US-10-465



Db 2626 AAGCTGGGCGTCTCCATTCAGAACCTGGTAAAGTCTACCGA---GATGGAGTGAAGGTG 2682  
Qy 495 GluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGly 514  
Db 2683 ---GCTGTCGATGGCTGGCACTGAATTTTATGAGGGCCAGATCACCTCTCTCTGGGC 2739  
Qy 515 HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuGlyCysGlyLeuCysProSer 534  
Db 2740 CACAATGGAGCGGGAGACGACCATGTCATCTGACCGGGTGTTCCTCCCGCAGCC 2799  
Qy 535 AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
Db 2800 TCGGGCACCGCTACATCTCTGGGAAAAGACATTCGC---TCTGAGATGACCCATC 2853  
Qy 555 ArgIleMetIleGlyLeuCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
Db 2854 CGCAGAACCTGGGGTCTCTCCAGCATAAAGTGTCTGTGACATGTGATCTGCGAA 2913  
Qy 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnIleIleGlnGlu 594  
Db 2914 GAACATCTGCTCTATGCCCGCTGAAAGGCTCTCTGAGNAGCACGCTGAAGCGGAG 2973  
Qy 595 ValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLys---AspAsnGlnAlaLys 613  
Db 2974 ATGGAGCAGATGCCCTGGATGTGTGTTGCCATCAAGCAAGCTGAAAGCAAAACAAGC 3033  
Qy 614 LysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnPro 633  
Db 3034 CAGCTGTGAGTGGATGAGAGAAAGCTATCTGTGGCTTGGCCCTTGTGCGGGATCT 3093  
Qy 634 LysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleVal 653  
Db 3094 AAGGTTGTCTATCTGATGAACCCACACAGTGTGTGGACCTTACTCCCGCAGGGGAATA 3153  
Qy 654 TrpAsn---LeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPhe 672  
Db 3154 TGGAGCTGCTGTGTAATACCGCAA---GGCCGACCATTTATCTCTACACACCAC 3210  
Qy 673 MetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLys 692  
Db 3211 ATGGATGAAGCGGACGCTCTGGGGACAGATTTGCCATCATCTCCATGGGAAGTGTGC 3270  
Qy 693 CysValGlySerSerMetPheLeuLysSerLysTyrGlyIleGlyTyrArgLeuSerMet 712  
Db 3271 TGTGTGGCTCTCTCCCTGTTCTGAAGAACCGATGGGACAGGCTACTACTGACCTTG 3330  
Qy 712 ----- 712  
Db 3331 GTCAAGAAAGATGTGAATCTCTCCCTCAGTTCTGTGCAGAAACAGTAGTAGTGTGTA 3390  
Qy 713 TyrIleAspLysTyrCysAla----- 719  
Db 3391 TACCTGAAAAGAGGAGACAGATGTTCTCAGAGCAGTTCTGATGCTGGCTGGCGCAGCGAC 3450  
Qy 720 -----ThrGluSerLysSerLeuValLysGlnHis 730  
Db 3451 CATGAGATGCACAGCTGACCATCGATGCTCTGTCTATCTCCAACTCATCAGGAGCAT 3510  
Qy 731 IleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPhe 750  
Db 3511 GTGTCTGAAGCCCGGCTGGTGAAGACATAGGGCATGAGCTGACCTATGTGTGTCATAT 3570  
Qy 751 LysAspMetAspLys-----PheSerGlyLeuPheSerAlaLeuAspSerHis---Ser 767  
Db 3571 GAAGCTGCTAAGAGGGAGGCTTTGTGGAACTCTTCATGAGATGATGACCGGCTCTCA 3630  
Qy 768 AsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLys 787  
Db 3631 GACCTGGCATTTCTAGTTATGGCATCTCAGAGACGACCTCGAAGAAATATTCCTCAAG 3690  
Qy 788 LeuGluValGluAlaGluIleAsp----- 795  
Db 3691 GTGGCCGAGAGAGTGGGGTGGATGCTGAGACCTCAGATGGTACCTTGGCCAGCAAGACA 3750

Qy 796 -----GlnAlaAspTyrSerValPheThrGln----- 804  
Db 3751 AACAGGGGGCTTTGGGGACAAGCAGAGCTGTCTTCGCCGCTTCACTGAAGATGATGCT 3810  
Qy 805 ---GlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu 823  
Db 3811 GCTGATCCAAATGATCTGACATAGACCCAGATCC-----AGAGAGACAGACTTG 3861  
Qy 824 LeuIleLeuSerGluThrLysAlaSerLeu---ValSerThrMetSerLeuTyrLysGln 842  
Db 3862 CTCAGTGGATGGATGGCAAGGGTCTCTACAGGTGAAAGGCTGGAACCTTACACAGCAA 3921  
Qy 843 GlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLysSerVal 862  
Db 3922 CAGTTTGGGCCCTTTTGTGGAGAGAGACTGCTAATTCAGACAGAGTTCGGAAGGATTT 3981  
Qy 863 ArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPheLeuVal 882  
Db 3982 TTTGCTCAGATTGCTTGGCAGCTGTCTTCTGTCATTGCCCTTGTGTTTCAGCTGATC 4041  
Qy 882 ----- 882  
Db 4042 GTGCCACCTTTTGGCAAGTATACCCAGCCTGGAACCTTCAGCCCTGGATGTACACGAACAG 4101  
Qy 882 ----- 882  
Db 4102 TACACATTTGTCAACATGATGCTCTCTGAGACACGGGAACCTTGAACCTTTAAACGCC 4161  
Qy 882 ----- 882  
Db 4162 CTCACCAAGACCCTGCTTGGGACCCGCTGTATGGAAGGAACCAATCCAGACACG 4221  
Qy 883 -----HisHisSerPheLysAsnAlaValProIleLysLeuValPro 897  
Db 4222 CCCTGCCAGGCGAGGAGAGAGTGGACCACTGCCCTGCCAGTTCC---CAGACCATCATG 4278  
Qy 898 AspLeuTyr----- 900  
Db 4279 GACCTCTTCAGAAATGGAACTGGACAATGCAGAACCTTCACTGTCATGCCAGTGTAGC 4338  
Qy 901 -----PheLeuLysProGly-----Asp 906  
Db 4339 AGCAGCAAAATCAAGAAGATGCTGCTGTGTGTCCTCCAGGGCAGGGGGCTGCTCCT 4398  
Qy 907 LysProHisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAspSerAspIleSer 926  
Db 4399 CCACAAAGAAACAAACAACTGCAGATATCTCTCAGGACCTGCAGAGGAACAACTTCG 4458  
Qy 927 Asp-----LeuIleSerPhePheThrSerGlnAsnIleMet 938  
Db 4459 GATTATCTGTGAAGACGTATGTGCAGATCATAGCCAAAGCTTAAAGAACAGATCTGG 4518  
Qy 939 ValThr----- 940  
Db 4519 GTGAATGAGTTAGTATGGCGCTTTTCCCTGGGTGTCACTAATACTCAAGCACTCCT 4578  
Qy 941 -----MetIleAsnAspSer-----AspTyrValSerValAlaPro 952  
Db 4579 CCGAGTCAAGAAGTTATGATGCCCAACAAATGAAGAAACACCTTAAAGCTGGCCAG 4638  
Qy 953 HisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrVal----- 967  
Db 4639 GACAGTTCTGCAGATCGATTTCTCAGAGCTTGGGAGATTTATGACAGAGCTGGACAC 4698  
Qy 968 -----PheAlaAlaValPheAsnSerThrMetValTyrSerLeuProIleLeuVal 984  
Db 4699 AGAAATAATGTCAAGGTGTGTTTCAATAACAGGGGTGGCATGCAATCAGCTCTTTCCTG 4758  
Qy 985 AsnIleIleSerAsnTyrTyrLeu----- 992  
Db 4759 AATGTATCAACATGCCATTTCTCCGGGCCCAACTGTCAAAGGAGAGAACCCCTAGCCAT 4818

```

Db      4820 TTCTTAGAGCTTTCTAAAGAACAG---GAAAGTAGGAAATTTTGATGAGAAATTTGATACA 4876
      1629 ThrLeuTrpTrpGlu 1633
      4877 ACAATGAGATGGAA 4891

RESULT 9
US-10-465-498-91
; Sequence 91, Application US/10465498
; GENERAL INFORMATION:
; APPLICANT: DENEFE, Patrice
; APPLICANT: ROSIER-MONTUS, Marie-Francoise
; APPLICANT: ARNOULD-REGUIGNE, Isabelle
; APPLICANT: FRADES, Catherine
; APPLICANT: LAUDIN, Laurent
; APPLICANT: LEMOINE, Cendrine
; APPLICANT: DUVERGER, Nicolas
; APPLICANT: ASSMANN, Gerd
; APPLICANT: RUST, Stephan
; APPLICANT: FUNK, Harald
; APPLICANT: BREWER, Bryan
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS CORRESPONDING TO HUMAN GENE ABC1
; FILE REFERENCE: ST99020USCNT1
; CURRENT APPLICATION NUMBER: US/10/465,498
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: FR99 07684
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: PCT/FR00/01595
; PRIOR FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: US10/018,714
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: US10/269,780
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 159
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 91
; LENGTH: 9497
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6765)..(6765)
; OTHER INFORMATION: n is a, c, g, or t
US-10-465-498-91

Alignment Scores:
Pred. No.:      8,86e-134      Length:      9497
Score:          1475.00      Matches:      490
Percent Similarity: 42.75%      Conservative: 318
Best Local Similarity: 25.93%      Mismatches: 624
Query Match:      17.51%      Indels:      458
DB:              6          Gaps:      65

US-10-090-458-5 (1-1642) x US-10-465-498-91 (1-9497)
QY      89 SerSerIleMetGlnLysValSerThrAspHisLeuProAspValIleIleThrGluGlu 108
Db      1435 TCTCCTCATGGAGTGTCAACTGACAAAGCTA---GAACCCATAGCAACAGATC 1491
QY      109 TyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLysProSerAsnPheValGly 128
Db      1492 TGGCTCATCAACAAGTCCATG-----GAGCTGTGGATGAGGAAAGTTCTGGGCTGGT 1545
QY      129 ValValPhe-----LysAspSerMetSerTyrGlu 138
Db      1546 ATGTGTTCACTGGAAATTAATCCAGGAGCATGAGTGGCCCATCATGTCAGTACAG 1605
QY      139 LeuArgPhePheProAspMetIle-----ProValSerSerIleTyrMetAsp 154
Db      1606 ATCCGAATGGACATTGCAATGTGGAGAGGACAAATAAAACCAAGGATGGGTACTGGGAC 1665
QY      155 -----SerArgAlaGlyCysSerLysSerCysGluAlaAlaGlnTyrTrpSerSerGly 172

```

```

Db      1666 CTGTGCTCTCAGCT-----GACCCCTTTGAGGACATCGCGTACGCTCTGGGGGGC 1716
QY      173 PheThrValLeuGlnAlaSerIleAspAlaIleIleGlnLeuLysThrAsnValSer 192
Db      1717 TTGGCTACTTTGAGGATGTGGTGAGCAGCAATCATCAGGGTGTCTGACGGGCACC-- 1773
QY      193 LeuTrpLysGluLeuGluSerThrLysAlaValIleMetGlyGluThrAlaValValGlu 212
Db      1774 -----GAGAAGAAAACCTGGTCTCTATATGCAACAGATG----- 1806
QY      213 IleAspThrPheProArgGlyValIleLeuIleTyrLeuValIleAlaPheSerProPhe 232
Db      1807 -----CCCTATCCCTGTGTACGTTGATGACATCTTTCTGCGGTGATGAGCGGTCAATG 1860
QY      233 GlyTyrPheLeuAlaIle-----HisIleVal 241
Db      1861 CCCCTCTTCATGACGCTGGCTGGATTTACTGAGTGGCTGTGATCATCAAGGGCATGTG 1920
QY      242 AlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThrAla 261
Db      1921 TATGAGAAGGAGGACCGCTGAAGAGACCATGGGATCATGGGCTGGACACAGCATC 1980
QY      262 PheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMet-----Ser 278
Db      1981 CTCTCGTTTACGCTGTTTCATT-----AGTAGCCCTCATCTCTTCTTGTGAGCGCTGGC 2034
QY      279 LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleVal 298
Db      2035 CTGCTAGTGGTCATCTCGAAGTTAGGAACCTGCTG---CCCTACAGTATCCCGAGGTG 2091
QY      299 IlePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThr 318
Db      2092 GTGTTTGTCTCTCGTGTGTTCTGTGTGTGACAAATCCTGCGAGTGTCTCTGATTAGC 2151
QY      319 ProLeuPheLysLysSerLysHisVal-----GlyIleValGluPhePheVal 334
Db      2152 ACATCTCTTCCAGACCAACCTGGCAGCGCTGTGGGGCAATCATCTACTTCCAGCTG 2211
QY      335 ThrValAlaPhe-----GlyPheIleGlyLeuMetIleLeu 347
Db      2212 TACCTGCCCTACGCTCTGTGTGGCATGGAGGACTACGTGGGCTTCCACATCAAGATC 2271
QY      348 IleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCysThrPhe 367
Db      2272 TTCGCTAGC-----CTGCTGTCTCTCT-----GTGGCTTTT 2301
QY      368 ValIleGlyIleAlaGlnValMethHisLeuGluAspPheAsnGluGlyAlaSerPheSer 387
Db      2302 GGGTTTGGCTGTGAGTACTTTGCCCTTTTGGAGAGCAGGCAATGGAGTGCAGTGGGAC 2361
QY      388 AsnLeuThrAlaGlyPro-----TyrProLeuIleIleThrIleIleMet 402
Db      2362 AACCTGTTTGGAGAGTCCCTGTGGAGGAAGATGGCTCAATCTCACCATCTCGTCTCCATG 2421
QY      403 LeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnValIlePro 422
Db      2422 ATCTGTTTGGACACCTTCTCTATGGGTGATGACCTGTATCATTTAGGCTGTCTTTCCA 2481
QY      423 GlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp----- 440
Db      2482 GGCCAGTACGGAATTCACAGGCCCTGGTATTTCTTCGACCAAGTCTCTACTGGTTTGGC 2541
QY      441 -----SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsn 454
Db      2542 GAGGAAAGTATGAGAGAGCCCTGGTTCCACACAGAGAGAAATATCAGAA----- 2595
QY      455 ValAsnGlyAsnIleSerPheSerGluIleIleGluProValSerSerGluPheValGly 474
Db      2596 -----ATCTGATGAGGAG-----GAACCCACCCACTTG----- 2625
QY      475 LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnVal 494

```





1235 ArgSerIleArgLysAspProPheArgAsnLeuSerThrLysSerIysAsnArgLys 1254  
 3697 AAATCAATGAGAAGGATCTCTTTCTTTAGAAATTTCTCCAAAGAGTAGTGTGTCTCAA 3756  
 1255 LeuProGluProPheAspAsnGluAspGluAspValLysAlaGluArgLeuLys 1274  
 3757 AATCCAGAGAACCA-----GAAGGAGAGATGAAGATGTTTCAGTGGAAAGATGAGA 3810  
 1275 ValLysGluLeuMetGlyCysGlnCysCysGluLysProSerIleMetValSerAsn 1294  
 3811 ACAGCAATGCTTGAATTTCTACTAATTTTGAAGAGGAGGATCATTCATTCAGGCTGT 3870  
 1295 LeuHisLysGluThrAspAspLysValPheLeuLeuSerArgLysValLysVal 1314  
 3871 CTACGAGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3930  
 1315 AlaThrLysTyrIleSerPheCysValLysLysGluLysLeuLysLeuLysPro 1334  
 3931 GCCACGAGAAATGCTCTCTCTGTTGTAGAAAGGTGAAGTTTATAGGATTTATAGGAC 3990  
 1335 AsnGlyAlaGlySerThrIleAsnIleValGlyAspIleGluProThrSer 1354  
 3991 AATGGAGCTGTAAGAGCATTCATTAAAGTGATTAAGTGGAGACACAAACCAATGCT 4050  
 1355 GlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAspSerLeuLys 1374  
 4051 GGACAAGTGCTACTG-----AAAGGAGCGGTGGAGGGGATGCCCTGGAG 4095  
 1375 CysMetGlyTyrCysProGlnIleAsnProLeuThrProAspThrThrLeuGlnGluHis 1394  
 4096 TTCCTGGGGTACTCCCTCAGGAGAACCGCTGTGCCCCCACTGACAGTGAGGAGCAC 4155  
 1395 PheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSer 1414  
 4156 CTGGAGGTGTACCGCGCGGTGAAGAGGCTGAGGAAGGGATGCTGAGGTGGCATCACA 4215  
 1415 ArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuPro 1434  
 4216 CGGTTAGTGGTGGCTCAAGCTCAGGACCGAGCTGGAAGTCTCCGGTGAAGACCTTGTCA 4275  
 1435 AlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThr 1454  
 4276 GAGGGAATTAACAGAAAGCTGTGTTTCTGAGCATCTGCGGAACCGCTCAGTGGTG 4335  
 1455 LeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAla 1474  
 4336 CTTCTGATGAGCGCTGACCGGATGAGCCCGAGGGGAGCAGCAGCAAAATGTGGCAGGC 4395  
 1475 IleArgThrAlaPheLysAsnArgLysAlaIleLeuThrHisTyrMetGlu 1494  
 4396 ATCCGGCCACCTTTAGAAACAGGAAGGGGTGCTCTTACACCCACTACATGGCA 4455  
 1495 GluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIle 1514  
 4456 GAGGCTGAGGCGGTGTGACCGAGTGGCCATCATGTGTATCTGGAGGTTGAGATGTATC 4515  
 1515 GlyThrValGlnHisLysLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeu 1534  
 4516 GTTCTCCCAACACCTGAAAGCAATTTGGCAAGATTAACCTGCTGGAGATGAAGTG 4575  
 1535 LysAspTrpIleGluAsnLeu---GluValAspArgLeuGlnArgGluIleGlnTyrIle 1553  
 4576 AAG-----AACCTGGCACAAGTGGAGCCCTCCATGTCAGAGATCTCGAGGCTT 4623  
 1554 PheProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAlaTyrLysIleProLys 1573  
 4624 TTCCTCCCGAGGTGTGCGGAGGAAGGTATCTCTCTGTATGTTTATAAGTTGCCAGTG 4683  
 1574 GluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluAlaLysHisAlaPhe 1593  
 4684 GAAGTGTGCAACCTTTAGCCCAAGCTTCTTCAATTAGAGAGGTTAAACAGAGCTTT 4743

1594 AlaIleGluGluTyrSerPheSerGluAlaThrLeuGluGlnValPheValGluLeuThr 1613  
 4744 GACCTAGAGGAGTACAGCCTCTCAGCTCTACCTCTGAGGAGGTTTCTCTGGAGCTTCC 4803  
 1614 LysGluGlnGlu-----GluGluAsp 1620  
 4804 AAGGAGCAGGAGCTGGGTGATTTTTCAGGAGGAT 4836

RESULT 8  
 US-10-796-280-369  
 ; Sequence 369, Application US/10796280  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CARGILL, Michele et al.  
 ; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
 ; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: CL001510  
 ; CURRENT APPLICATION NUMBER: US/10/796,280  
 ; NUMBER OF SEQ ID NOS: 68533  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 369  
 ; LENGTH: 5185  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-796-280-369

Alignment Scores:  
 Pred. No.: 1,876-281 Length: 5185  
 Score: 2967.50 Matches: 664  
 Percent Similarity: 59.22% Conservative: 322  
 Best Local Similarity: 39.88% Mismatches: 548  
 Query Match: 35.22% Indels: 131  
 DB: Gaps: 34

US-10-090-458-5 (1-1642) x US-10-796-280-369 (1-5185)

QY 5 IleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLysAsnTyrIleuile 24  
 DB 182 ATGAACAGAAAGCGTGTATCAGCAACCAAGCAGCTCTGTGCAAGATTTCTTAAG 241  
 QY 25 LysCysArgThrLysLysSerValGlnGluLeuPheProLeuPhePheLeuPhe 44  
 DB 242 AAATGGAGGATGAAAGAGAGAGCTTATTGGAATGGGCGCTCTCAATACTCTTAGGACTG 301  
 QY 45 TrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGluValProAsnIle 64  
 DB 302 TGATTTGTCTGTTTTCAGATTCCATGAGAAATGTCCAGTTTCTGGAAATGGCTCCTCAG 361  
 QY 65 GluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGlyTyrThrProVal 84  
 DB 362 AATCTGGGAGGGTAGATAAATTTAATAGCTCTCTTTAATGGTTGTGTATACCAATA 421  
 QY 85 ThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu-----Pro 101  
 DB 422 TCTAATTTAAACCCAGCAGATTAATGAATAAACACAGCAGCTTGCTCTCTTTTGAAGGAACA 481  
 QY 102 AspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeuSer 121  
 DB 482 AGTGTCATTTGGGCGCCAAATAAACAACACAGCAGCAATACCTCTCGAAATTTACCA 541  
 QY 122 LysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArgPhe 141  
 DB 542 TATGCT-----ATGGGAATCATCTTTAATGAACCTTCTCTTATAAGTTAATATTT 592  
 QY 142 PheProAspMetIleProValSerSerIleTyrMet---AspSerArgAlaGlyCys--- 159  
 DB 593 TTCAGGAGATAT-----AACAGTCCACTTTGGAAGAAGATTTCTCAGCTCATGTGG 646  
 QY 160 -----SerLysSerCysGluAlaIleGlnTyrTrpSerSerGlyPheThrVal 175  
 DB 647 GATGGATATGTGAGTGTTCATGATGACCAATATGCAATAGAGGATTTGTGGCT 706  
 QY 176 LeuGlnAlaSerIleAspAlaIleIleGlnLysThrAsnValSerLeuTrpLys 195

Db 1654 ATTTACGAAGCCAAATCACTGCAATACCTTGGTCACAGTGGAGCTGCAAGAGTCAACACATG 1713  
Qy MetAsnIleLeuCySgLYLeuCySProProSerAspGlyPheAlaSerIleTyrGlyHis 543  
Db 1714 CTAACATCTTCTAGTGGTTGTCTGCTCCACCACCAAGGTTCACTCAACCTATATAACAAT 1773  
Qy ArgValSerGluIleAspGluMetPheGluAlaArgYsMetIleGlyIleCysProGln 563  
Db 1774 AAGCTTTCAAGAAATGGCTGACCTAGAAATCTCAGCAAGCTGACCGAGTTGTGCACAA 1833  
Qy LeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIle 583  
Db 1834 TCCAATGTGCAATTTGACTCTCTCACTGTAAAGAGAAACCTCAGACTCTTTGCTAAATA 1893  
Qy LysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAsp 603  
Db 1894 AAAGGGATT-----CTGCCACAAGAAGTGGATAAA----- 1923  
Qy MetGlnThrIleLysAspAsnGlnAlaLysLeuSerGlyGlyGlnLysArgLysLeu 623  
Db 1923 ----- 1923  
Qy SerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAla 643  
Db 1924 -----GAGATTTTCCTGTTGGATGAACCACTGCT 1953  
Qy GlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsn 663  
Db 1954 GGATGGATCCCTTTCAAGACCAAGATGGACCTCTGAAAGAACGCAAAACAGAC 2013  
Qy ArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLys 683  
Db 2014 CGCTGATCTCTTCAGTACCACTTCATGGATGAGCCGACATCTCGCGGACAGGAAA 2073  
Qy AlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLys 703  
Db 2074 GTATTTCTCTCCCAAGGAAGCTAAAGTGGCGGCTCTCTTGTCTTAAGAGAAA 2133  
Qy TrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeu 723  
Db 2134 TGGGGGATTGGATATCACTTAAGCTTGCATTAATGAAATATGTGTGAGGAAAAATA 2193  
Qy SerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGln 743  
Db 2194 ACATCACTTGTAAACAGCACATCCCTGTATGCCAAATATATCAGCCAAAGCGAAGGAAA 2253  
Qy LeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeu 763  
Db 2254 CTTATTATACATTACCTTAGAAGAACCAAAATAATTCAGAACTTTTACAAGGATCTT 2313  
Qy AspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrLeuGluAsp 783  
Db 2314 GATAGCTATCTGACCTAGGAATTTGAGAATTTATGTTTCCATGACAACTTTGAATGAA 2373  
Qy ValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThr 803  
Db 2374 GTATTCCTGAAGCTAGAGGAAAATCTACAAATTAATGAATCGGACATTTGCTATTTGGGA 2433  
Qy GlnGlnProLeuGluGluLysMetAspSerLysSerPheAspGluMetGlnGlnSerLeu 823  
Db 2434 GAAGTACAAAGCGGAAAAGCTGACGACACTGAAAGGCTTGTGTGATGGAACAAAGTCCCT 2493  
Qy LeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyrLysGlnGln 843  
Db 2494 TCTTCACCTTAACAGATGAGAAAGACA---ATAGGTGGTGGCTCTCTGGGACAGCAA 2550  
Qy MetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLysSerValArg 863  
Db 2551 ATCTCGCAATTCGAAGGGTTTCGCTGTGTTAAAGTAAAGCATGAAGAAAGATCTTTTA 2610  
Qy SerValLeuLeuLeuIlePhePheThrValGlnIlePheMetPheLeuValHis 883  
Db 883 -----

Db 2611 GCACGTCTATTAAATCTAATGGCTGGATTTTGGCCCTCTTCTTGTGGAGTATACCATGGTG 2670  
Qy HisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLys 903  
Db 2671 AAAATATATCAAAACAGATTACACC---TGGGAATTTCTCTCTCATTTGATTTTCTTGCT 2727  
Qy ProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAspSer 923  
Db 2728 CTTGGACAACAACACACATGACCTCTCACTCAACTACTGATCATCAATAAACAGGCGCA 2787  
Qy AspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThrMet----- 941  
Db 2788 AGCATTTGATGACTTTATACAGTCTGTGGAGCACGACCATGCTTTTGAAGTGGATGCA 2847  
Qy -----IleAsnAspSerAspTyrValSerValAlaProHisSerAlaLeuAsn 958  
Db 2848 TTTGGAACCTAGAAATGGCACAGATGACCCATCT-----TATAATGGAGCCATCACA 2898  
Qy ValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyr 978  
Db 2899 GTGTGTTGTAATGAAAGAAATTCAGCTTTTCGTAGCATGCAATGCCAAAGAAATGAAT 2958  
Qy SerLeuProIleLeuValAsnIleSerAsnTyrTyrLeuTyrHisLeuAsnValThr 998  
Db 2959 TGCITCCAGTCTTATGGACATTTGATAGTAAATGGCTACTTGGATGGTTAAACCATCA 3018  
Qy GluThrIleGlnIleTrpSerThrPhePheGlnGluIleThrAspIleValPheLys 1018  
Db 3019 GTACATATCCGAACCTGAAAGAGATACATTTTGGAGAAATGGACAGCAATCCCAATCGGA 3078  
Qy IleGluLeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProProTyr 1038  
Db 3079 TTCTGGCATATATCATGTTCTGCTG-----GTTTTAATCATCGAGTCCCACTTAC 3132  
Qy PheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSer 1058  
Db 3133 ATTGCCATGACGAGCATCGATGATTATAAGAACAGAGCTCGTCCCACTACGAGTTCC 3192  
Qy GlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePhe 1078  
Db 3193 GGACTCTCCCTTCTGCTTACTGTTGGGACGCGCTGGTGGATGTTCCTCTACTTTC 3252  
Qy IleIleLeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyr 1098  
Db 3253 TTGTGCTCTTCTTTATATATTTAATGAGCTACATTTCAAACTTCGAAGACATGCTACTT 3312  
Qy ThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeu 1118  
Db 3313 ACAATAATTCATATATTTCAAATCCCATGCTGCTGTGTGTTATTCCTTTCCCTCATCTTC 3372  
Qy PheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSer 1138  
Db 3373 ATGACATAGTGATTTCTTCTCATCTTCGCAAGGGGAGAAAAATAGTGGCATTTGGTCA 3432  
Qy PheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMet 1158  
Db 3433 TTTTCTTTCTATGTTGTCACTGATTCCTGTGGCTGGATTTGGCTTCAGTATCTTTC--- 3489  
Qy GlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrPro 1178  
Db 3490 -----GAAAGTGATATTCATTTTATTTTAAATACCACCTGCCACA 3537  
Qy LeuLeuGlyCysLeuIle-----SerPheIleLysIleSer-----TrpLysAsnVal 1194  
Db 3538 ATGATTTGGCTGTTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3597  
Qy ArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSerPro 1214  
Db 3598 CGAATGGATGACACCCATTT-----CTGGTATTCTTAATTCCT 3636  
Qy TyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysTyrGlyGly 1234  
Db 3637 TTCTCTTCAATTTTATCATTTTCTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3696

Db 4684 GAAGATGTCACACCTTTAGCCCAAGCTTTCTTCAAAATTAGAGAAGGTTAAACAGAGCTTT 4743  
 Qy 1594 AlaileGluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThr 1613  
 Db 4744 GAACATAGAGGATACAGCCCTCACAGCTACCTGAGCAGGCTTTTCTGGAGCTCTCC 4803  
 Qy 1614 LysGluGlnGlu-----GluGluAsp 1620  
 Db 4804 AAGGAGCAGGAGCTGGGTGATTTTGAGGAGGAT 4836  
 RESULT 7  
 US-10-788-792-45  
 ; Sequence 45, Application US/10788792  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bayer Pharmaceuticals Corporation  
 ; APPLICANT: Eveleigh, Deepa  
 ; APPLICANT: Bigwood, Douglas  
 ; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE  
 ; FILE REFERENCE: 5152  
 ; CURRENT APPLICATION NUMBER: US/10788,792  
 ; CURRENT FILING DATE: 2004-02-27  
 ; PRIOR APPLICATION NUMBER: US 60/450,855  
 ; PRIOR FILING DATE: 2003-02-28  
 ; NUMBER OF SEQ ID NOS: 254  
 ; SOFTWARE: Patent in version 3.2  
 ; SEQ ID NO 45  
 ; LENGTH: 5677  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-788-792-45  
 Alignment Scores:  
 Pred. No.: 3e-295 Length: 5677  
 Score: 3108.00 Matches: 672  
 Percent Similarity: 60.02% Conservative: 319  
 Best Local Similarity: 40.70% Mismatches: 536  
 Query Match: 36.89% Indels: 124  
 DB: 6 Gaps: 26  
 US-10-090-458-5 (1-1642) x US-10-788-792-45 (1-5677)  
 Qy 6 ArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLysAsnTyrLeuLeuLys 25  
 Db 148 AGAAGATCAGTGTGTCAACAACTTGGGCTTATTATGCAAGAACTTTCTTAAAAA 207  
 Qy 26 CysArgThrLysSerSerValGlnGluLeuPheProLeuPheLeuPheLeuPheTrp 45  
 Db 208 TGGAGATGAAGAGAGCTTAAATGAATGCGTGAATTCATTCCTCTACTACTTTGT 267  
 Qy 46 LeuLeuLeuLeuSerMetMetHisProAsnLysLysTyrGluGluValProAsnLeuGlu 65  
 Db 268 TTGTATATATATCTCATATGTCATCAAGTAAATGATTTTCTTCACTGCTTACCAGGAC 327  
 Qy 66 LeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGlyTyrThrProValThr 85  
 Db 328 CTGGAGCGGTAGATACATTAAATGAATCCAGATTTCTGTGTATACACACCTGTCACC 387  
 Qy 86 AsnIleThrSerSerIleMetGlnLysValSerThrAsp----- 98  
 Db 388 AACACGACCAACAGATATATGAATAAGTAGCTCTACTCCCTTCTGCGAGGTAAGAG 447  
 Qy 99 -----HisLeuProAspValIleThrGluGluTyrThrAsnGluLysGluMetLeu 116  
 Db 448 GTCTGGGAGTGCAGATGAGGAAGATATTAAAGAAATTCACA----- 489  
 Qy 117 ThrSerSerLeuSerLysProSerAsnPheValGlyValValPheLysAspSerMetSer 136  
 Db 490 -----GCAATATTCTGAGAAATAGTAGAGTCACTTTACTTAATCACTACTCA 540  
 Qy 137 TyrGluLeuArgPhe---PheProAspMetIleProValSerSerIleTyrMetAspSer 155  
 Db 541 TATCAITTTGAAGTCTTCTGTAGGACATGGAATGCCAGCAAGAGGACCAAGGACCAT 600

Qy 156 ArgAlaGlyCysSerLysSer-----CysGluAlaAlaGlnTyrTrpSer 170  
 Db 601 ACAGCTCATTTGTTATGAACAACAAATGAAGATGTTTACTGTGAGTTCAGTATTGTGAAG 660  
 Qy 171 SerGlyPheThrValLeuGlnAlaSerIleAspAlaAlaIleIleGlnLeuLysThrAsn 190  
 Db 661 GAAGGTTTTGTGGCTCTTCAAGCTGCCATTAATCTCTATTATAGAATCACAACAAT 720  
 Qy 191 ValSerLeuTrpLysGluLeuGluSerThrLysAlaValIleMet----- 205  
 Db 721 CACTCAGTATGGAGGAGCTGATGTCTACTGGAATAATATGAAGATGCATCTCTTC 780  
 Qy 206 ---GlyGluThrAlaValValGluIleAspThrPheProArgGlyValIleLeuLeuTyr 224  
 Db 781 ATTGTCATCAATCAGGAGTTATA-----ACTGATTGTACTCTTTTTC 822  
 Qy 225 LeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIleValAlaGluLys 244  
 Db 823 TGCATTATTCT 882  
 Qy 245 GluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThrAlaPheTrpLeu 264  
 Db 883 ---AAAAGATGAAGGCTTGATGATCAATGATGGTCTTCGGGATTCAGCGTTCTGGCTC 939  
 Qy 265 SerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuMetAlaValIle 284  
 Db 940 TCCTGGGTTTGTCT 999  
 Qy 285 AlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePheLeuLeuPhePhe 304  
 Db 1000 ATAAGATCTACCCAGTTTATCATTTTGTCTGGCTTCATGTGATGCTTCAGCTCTTCTC 1059  
 Qy 305 LeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeuPheLysLysSer 324  
 Db 1060 CTGTATGATTTATCTTTGGTAGCTTTGGCTTTCTTAATAGAGCATCTTGGTAAAGAAATCT 1119  
 Qy 325 LysHisValGlyIleValGluPhePheValThrValAlaPheGlyPheIleGlyLeuMet 344  
 Db 1120 TTCTCACCGGCTTGGTGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1176  
 Qy 345 IleIleLeuLeuGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHis 364  
 Db 1177 ACATCAGTGTACAGACACCTTCTCTGATCTCTTGGAGTGGATTTTAAGCTTGTCTAGTCC 1236  
 Qy 365 CysThrPheValIleGlyIleAlaGlnValMetHisLeuGlu---AspPheAsnGluGly 383  
 Db 1237 TTGCTTCTAGCTTGAATGGCCAGCTTTTACACTGGACTATGATTTGAATCTTAAT 1296  
 Qy 384 AlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIleMetLeu 403  
 Db 1297 GCA---TTTCTCATCCATCGACGGCTCAAACTCTCATTTAGCAACAAATTTTCATGTTG 1353  
 Qy 404 ThrLeuAsnSerIlePheTyrValLeuAlaValTyrLeuAspGlnValIleProGly 423  
 Db 1354 GCATTGACATTCCTCTATCTGGCATTTGGCATTTTACTTTGAAAAAATTTTGCCAAAT 1413  
 Qy 424 GluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrpSerLysSer 443  
 Db 1414 GAATATGGACATCACCTCCACCTTTGTTTCTGAGTCCCTCATTTTGTCTCAACA 1473  
 Qy 444 LysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGlu 463  
 Db 1474 CAAAGAGCTGATCAGTGGGCTTGAAGATGAATGGATGCGCATCTCTCTCTCTCTCTCT 1533  
 Qy 464 IleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIleSerGlyIle 483  
 Db 1534 TCTTTTGAACAGGCTCCAGATTTCAAGGGAAGAGCCATCAGATCAGAAATGTT 1593  
 Qy 484 GlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAsp 503  
 Db 1594 ACAAAGAATATAAGGAAGCCCTGATAAAATAGAGCTTGAAGATCTGGTATTGAC 1653  
 Qy 504 IleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeu 523





541 TATCAATTGAAGTCTTCTGTAGACATGGAATGCCAGCAAGAAAGAGAGCAAGACCAT 600  
156 ArgAlaGlyCysSerLysSer-----CysAlaAlaAlaGlnTyrTrpSer 170  
601 ACAGCTCATTTGTTATGAACAATAAGAGATGTTTACTGTGAAGTTTTCATATTGGAAG 660  
171 SerGlyPheThrValLeuGlnAlaSerIleAspAlaAlaIleIleGlnLeuLysThrAsn 190  
661 GAAGGTTTGTGCTCTTCAAGTCGCCAATTAATGCTGCTATTATAGAAATCACAAAT 720  
191 ValSerLeuTrpLysGluLeuGluSerThrLysAlaValIleMet----- 205  
721 CACTCAGTAGGAGAGCTGATGTCAGTTACTGGAAAAAATATGAAGATGCTCTTC 780  
206 ---GlyGluThrAlaValValGluIleAspThrPheProArgGlyValIleLeuIleTyr 224  
781 ATTGGTCAATCAGGAGTTATA-----ACTGATTTGTACCTTTTTC 822  
225 LeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIleValAlaGlnLys 244  
823 TGCATTATTTCATTTTCTCCATTCATTACTATGATCTGTTAATGTCAAGAGAGAGG 882  
245 GluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThrAlaPheTrpLeu 264  
883 ---AAAAGGATGAAGCCTTGATGACATGATGGTCTTCGGGATTCAGCGTTCGGCTC 939  
265 SerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeuMetAlaValIle 284  
940 TCTCGGGTTGCTCATGCTGCTGCTTTCATCTTCATTATGCGCTTCTTCTGGCATTGTT 999  
285 AlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePheLeuLeuPhePhe 304  
1000 ATAAAGTCTACCCAGTTTATCATTTTGTCTGGCTTCATGCTGATGCTTCAGCTCTTCTC 1059  
305 LeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeuPheLysSer 324  
1060 CTGTATGGATTATCTTTGTGTAGCTTTGGCTTTCTTAATGAGCATCTTGTGAAGAATCT 1119  
325 LysHisValGlyIleValGluPhePheValThrValAlaPheGlyPheIleGlyLeuMet 344  
1120 TTCTCACCGGCTGTGTGTCTCTCTCCCTGCTGCTTTTGGGGTGTCTGGG---TTC 1176  
345 IleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHis 364  
1177 ACATCACTGTACAGACACCTTCTGTCATCTTGGAGTGATTTAAGCTTGTAGTCC 1236  
365 CysThrPheValIleGlyIleAlaGlnValMetHisLeuGlu---AspPheAsnGluGly 383  
1237 TTTGCTTTCATGCTTGGATGGCCAGCTTTTACACTTGGACTGATTTGAATTTCTAAT 1296  
384 AlaserPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIleIleMetLeu 403  
1297 GCA---TTTCTCATCCATCGGACGGCTCAAAATCTCATTTAGCAACAAATTTTCATGTTG 1353  
404 ThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnValIleProGly 423  
1354 GCATTTGACACTGCTCATCTGCGATTGGCGATTACTTTGAAATAATTTTGCCAAAT 1413  
424 GluPheGlyLeuArgSerSerLeuTyrPheLeuLeuProSerTyrTrpSerLysSer 443  
1414 GAATATGGACATCGACCTCCACCTTTGTTTTCCTGAAGTCTCATTTTGGTCTCAACA 1473  
444 LysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGlu 463  
1474 CAAAGACTGATCACGTGGCCCTTGAAGATGAATGATGCGATGCGCTTTCATTTTCATGAC 1533  
464 IleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIleSerGlyIle 483  
1534 TCTTTTGACACGGCTCCAGNATTCGAAGGAAGAAGAGCCATCAGATCAGAAATGTT 1593  
484 GlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAsp 503  
1594 ACAAAAAGATATAAAGGAAGCCTGTATAAATAGAAGCCTTGAAGATCTGTATTGAC 1653

504 IleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeu 523  
1654 ATTTACGAGGCCAAATCATCTGCATATCTTGGTCAAGTGGAGCTGGAAGTCAACACTG 1713  
524 MetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIleTyrGlyHis 543  
1714 CTAACATCTTCTAGTGGTGTCTGTCTCCACCAAGAGTTCAGTCACCATCTATAACAT 1773  
544 ArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIleCysProGln 563  
1774 AAGCTTTCAGAAATGGCTGACCTAGAAAATCTCAGCAAGCTCAGCGAGTTGTGCCAA 1833  
564 LeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIle 583  
1834 TCCAATGTCAATTTGACTTCTCTCAGTGAAGAGAAAACCTCAGACTCTTTTGTCTAAATA 1893  
584 LysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAsp 603  
1894 AAAGGGATT-----CTGCCACAAGAAGTGGATAA----- 1923  
604 MetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLysArgLysLeu 623  
1923 ----- 1923  
624 SerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAla 643  
1924 -----GAGATTTTCTGTGGATGAACCACTGCT 1953  
644 GlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsn 663  
1954 GGATTTGATCCCTTTTCAAGACACCAAGTATGGAACCTTCTGAAGAACGCAACAGAC 2013  
664 ArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLys 683  
2014 CGCTGTATCTCTTCTCAGTACCCAGTTCATGCGATGAGCCGACATCTCTGGCGGACAGAAA 2073  
684 AlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLys 703  
2074 GTATTTCTCTCCCAAGGAAGCTAAAGTGGCGGGCTCTCTTTGTTCTTAAAGAGAAA 2133  
704 TrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeu 723  
2134 TGGGGATGATGATATCATCTTAAGCTTGAAGTAAATGAATATCTGTGTAGGAAAAACATA 2193  
724 SerSerLeuValLysGlnHisIleProGlyValThrLeuGlnGlnAsnAspGlnGln 743  
2194 ACATCACTTGTAAACAGCACATCCCTGATGCCAAATTTATCAGCCAAAACGAGGAAAA 2253  
744 LeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeu 763  
2254 CTTATTTATACATTACCTTTAGAAAGAACAAATAAATTTCCAGAACCTTACAGGATCTT 2313  
764 AspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAsp 783  
2314 GATAGCTATCTGACCTAGGAATTTAGATTTATGGTGTTTTCCATGACAACTTTGAATGAA 2373  
784 ValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThr 803  
2374 GTATTTCTGAAGCTAGAGGAAATCTACATTAATGAATCGACATTTGCTATTTGGGA 2433  
804 GlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu 823  
2434 GAAGTACAAGCGGAAAAAGCTGACGACACTGAAAGGCTTTGTGAGATGGAAACAGTCTC 2493  
824 LeuIleLeuSerGlnThrLysAlaSerLeuValSerThrMetSerLeuTyrLysGlnGln 843  
2494 TCTTCATCTTAACAGATGAGAAAGACA---ATAGGTGGTGTGGCTCTCTGGCCACAGCAA 2550  
844 MetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLysValArg 863  
2551 ATCTGCGCAATTGCAAGGGTTCGTTGTTAAAGTTAAAGCATGAAGAAAAGCTCTTTTA 2610



Db 1594 ACAAAGATATAAAGGAAAGCCTGATAAAATAGAAAGCCTTGAAAGATCTGGTATTGGAC 1653  
Qy IleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeu 523  
Db 1654 ATTACAGAGCCAAATCACTGCAATACCTTGGTCAGACAGCTGGAAGTCAACACATG 1713  
Qy MetAsnIleLeuGlyGlyLeuGlyCysProProSerAspGlyPheAlaSerIleTyrGlyHis 543  
Db 1714 CTAACATCTTAGTGGGTGTCTGTTCCCAACCAAGGTTTCAGTCACCATCTATAACAA 1773  
Qy ArgValSerGluIleLeuLeuMetPheGluAlaArgLysMetIleGlyLeuGlyCysProGln 563  
Db 1774 AAGCTTTTCAGAAATGGCTGACCTAGCAAAATCTCAGCAAGCTGACCGAGTTTGTCCACAA 1833  
Qy LeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIle 583  
Db 1834 TCCAATGTGCAATTTGACTCTCTCCTCAGTGAAGAGAAACCTCAGACTCTTGTGTAATA 1893  
Qy LysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAsp 603  
Db 1894 AAAGGATT-----CTGCCAAGAAAGTGGATAAA----- 1923  
Qy MetGlnThrIleLysAspAsnGlnAlaLysLeuSerGlyGlyGlnLysArgLysLeu 623  
Db 1923 ----- 1923  
Qy SerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAla 643  
Db 1924 -----GAGATTTTCTCTGTGGATGAACCAACTGCT 1953  
Qy GlyMetAspProCysSerArgHisIleValTyrAsnLeuLeuLysTyrArgLysAlaAsn 663  
Db 1954 GGATGGATCCCTTTCAAGACACCAAGTATGAACCTTCTGAAGAACCAAAACAGAC 2013  
Qy ArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLys 683  
Db 2014 CGCGTGATCCTCTTCAGTACCCAGTTCATGGATGAGCCGACATCTCGCGGACAGGAAA 2073  
Qy AlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLys 703  
Db 2074 GTATTTCTCCCAAGGAAAGCTAAAGTGCAGCGGCTCTCTTTGTTCTTAAAGAGAAA 2133  
Qy TrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeu 723  
Db 2134 TGGGGGATTGGATATCATCTTAAGCTTGCAGTTAAATGAATATGTCTGAGGAAACATA 2193  
Qy SerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGln 743  
Db 2194 ACATCACTGTGTTAAACAGCACATCCCTCGATGCCAAATTTACGCCAAAAGCGAAGAAA 2253  
Qy LeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeu 763  
Db 2254 CTATTTATACATTCACCTTAGAAGAACAAATAAATTTCCAGAACTTTACAGGATCTT 2313  
Qy AspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAsp 783  
Db 2314 GATAGCTATCCTGACCTAGGAATGGAATATGTTGTTCCATGACCACTTTGGAATGAA 2373  
Qy ValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThr 803  
Db 2374 GTATTCCTGAGCTAGAGAGAAATCTACAATTAATGAATCGACATTCGCTATTTTGGGA 2433  
Qy GlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu 823  
Db 2434 GAAAGTCAAGCGGAAAGAGTGCAGACACTGAAAGGCTTTGTCAGATGGAACAAAGTCTC 2493  
Qy LeuIleLeuSerGlnThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGln 843  
Db 2494 TCTTCACTAACAGATGAGAAAGACA---ATAGGTGGTGTGGCTCTCTGGCCACAGCAA 2550  
Qy MetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLysSerValArg 863  
Db 2551 ATCTCGCAATTCGAAGGGTCTGTTGTTAAAGTAAAGCATGAAAGAAAGCTCTTTTA 2610

Qy 864 SerValLeuLeuLeuLeuLeuPhePheThrValGlnIlePheMetPheLeuValHis 883  
Db 2611 GCATGCTATTAAATTAATGCTGGATTTGGCCCTCTCTTGTGGAGTATACCATGCTG 2670  
Qy HisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLys 903  
Db 2671 AAAATATATCAAAACAGTTACACC---TGGGAACCTTCTCTCTCATTTGTATTCTTCT 2727  
Qy ProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAspSer 923  
Db 2728 CTGGACCAACCAACCATGACCCCTCTCACTCACTACTGATCATCAATAAACAAGGGGCA 2787  
Qy AspIleSerAspLeuLeuSerPhePheThrSerGlnAsnIleMetValThrMet----- 941  
Db 2788 AGCATTGATGACTTTATACAGTCTGTGAGCACCAGAACATAGCTTTAGAAAGTGGATGCA 2847  
Qy -----IleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsn 958  
Db 2848 TTTGGAACCTAGAAATGGCAGAGTACCCATCT-----TATAATGGAGCCATCATCA 2898  
Qy ValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyr 978  
Db 2899 GTGTGTTGTAATGAAAGAAATACAGCTTTTCGTAGCATGCAATGCCAAAGATTTGAAT 2958  
Qy SerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThr 998  
Db 2959 TGTCTCCAGTTCTTATGACATTTGTAGTAATGGGCTACTTGAATGGTTTAAACCATCA 3018  
Qy GluThrIleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLys 1018  
Db 3019 GTACATATCCGAACCTGAAAGAAAGTACATTTTGGAGAAATGGACAGCAATCCCAATCGGA 3078  
Qy IleGluLeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProProTyr 1038  
Db 3079 TTCTCTGGCATATATCATGTTCTGCGTG-----GTTTTAACATCGAGTGGCCACCTTAC 3132  
Qy PheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSer 1058  
Db 3133 ATTGCAATGACGACATCGATATTATGAACAGAGCTCGTCCGACCTACGAGATTTC 3192  
Qy GlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePhe 1078  
Db 3193 GGACTCTCCCTCTCTGCTTACTGTTGGGCGAGCGCTGGTGGATGTTTCCCTGTACTTC 3252  
Qy IleIleLeuLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyr 1098  
Db 3253 TTGGTCTTCGTTTATATATATTAAATGAGCTPACATTTCAAACCTTCGAAGACATGCTACTT 3312  
Qy ThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeu 1118  
Db 3313 ACAATAATTCATATTATTAATCCCATGCCATGCTGCTGTTGGTATTCTCTTTCCTCATCTC 3372  
Qy PheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSer 1138  
Db 3373 ATGACATACGTGATTTCCCTTCATCTTTTCGAAGGGGAGAAAAATAGTGGCATTTGTGCA 3432  
Qy PheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMet 1158  
Db 3433 TTTTGTCTCATGTTGTCACGATTTCTCTGAGCTGGATTTGGCTTCAGTATCTTC--- 3489  
Qy GlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrPro 1178  
Db 3490 -----GAAAGTGATATTCCATTTATCTTCTCTCTTTTAAATACCACCTGCCACA 3537  
Qy LeuLeuGlyCysLeuIle-----SerPheIleLysIleSer-----TrpLysAsnVal 1194  
Db 3538 ATGATGGCTGTTGTTCTTATCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3597  
Qy ArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSerPro 1214  
Db 3598 CGAATGGATGACAGCCATTT-----CTGGTATTCTCTAAATCTCT 3636

1574 GluAspValGlnSerLeuSerGlnSerPhePheLeuGluGluAlaLysHisAlaPhe 1593  
 4684 GAAGATGTCACACCTTTAGCCCAAGCTTCTTCAATATAGACAGAGCTTT 4743  
 1594 AlaileGluGluTySerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThr 1613  
 4744 GACCTAGAGAGTACAGCCTCTCACAGTCTACCCCTGGAGCAGGTTTCTCGAGCTCTCC 4803  
 1614 LysGluGlnGlu-----GluGluAsp 1620  
 4804 AAGGACGAGGCTGGTGATTTTGAGGAGGAT 4836

RESULT 5

PCT-US04-07268-45  
 ; Sequence 45, Application PC/TUS0407268  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bayer Pharmaceuticals Corporation  
 ; APPLICANT: Eveleigh, Deepa  
 ; APPLICANT: Bigwood, Douglas  
 ; TITLE OF INVENTION: EXPRESSION PROFILES FOR BREAST CANCER AND METHODS OF USE  
 ; FILE REFERENCE: 5152  
 ; CURRENT APPLICATION NUMBER: PCT/US04/07268  
 ; CURRENT FILING DATE: 2004-02-27  
 ; PRIOR APPLICATION NUMBER: US 60/450,655  
 ; PRIOR FILING DATE: 2003-02-28  
 ; NUMBER OF SEQ ID NOS: 254  
 ; SOFTWARE: PatentIn version 3.2  
 ; SEQ ID NO 45  
 ; LENGTH: 5677  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 PCT-US04-07268-45

Alignment Scores:  
 Pred. No.: 3e-295 Length: 5677  
 Score: 3108.00 Matches: 672  
 Percent Similarity: 60.02% Conservative: 319  
 Best Local Similarity: 40.70% Mismatches: 536  
 Query Match: 36.89% Indels: 124  
 DB: 1 Gaps: 26

US-10-090-458-5 (1-1642) x PCT-US04-07268-45 (1-5677)

6 ArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLysAsnTrpLeuLeuLys 25  
 148 AGAAGATCATGCTGTGTCAACAACTTGGGCTTATTATGCAAGAACTTTCTTAAAAA 207  
 26 CysArgThrLysLysSerSerValGlnGluLeuLeuPheProLeuPhePheLeuPheTrp 45  
 208 TGGAGAATGAAGAGAGCTTATGGAATGGCTGAATCATTGCTCTACTACTTTGT 267  
 46 LeuileLeuileSerMetMetHisProAsnLysLysTrpGluGluValProAsnileGlu 65  
 268 TTGTATATATATCTCATATGTCATCAAGTAAATGATTTTCTTCACTGCTTACCAGGAC 327  
 66 LeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLeuGlyTrpThrProValThr 85  
 328 CTGGACGGGTAGATACATTAATGAATCAAGATTTCTGTTGATACACCTGTCACC 387  
 86 AsnileThrSerSerileMetGlnLysValSerThrAsp----- 98  
 388 AACACGACCCACACAGATAATGAATAAAGTAGGCTCTCTCCCTCTCTGGCAGGTAAGAG 447  
 99 -----HisLeuProAspValileleThrGluGluTyThrAsnGluLysGluMetLeu 116  
 448 GTCTTGGGCTGCCAGATGAGGAAGTATTAAAGAAATTCACA----- 489  
 117 ThrSerSerLeuSerLysProSerAsnPheValGlyValValPheLysAspSerMetSer 136  
 490 -----CCAAATATCTCGAAGAAATAGTAAGAGTACCTTTACTTAATACATCTCA 540  
 137 TyrGluLeuArgPhe---PheProAspMetileProValSerSerileTrpMetaspSer 155

541 TATCATTTGAAGTCTTGTGTAGACATGGAATGCCAGAAAGAGAGACAGGACCAT 600  
 156 ArgAlaGlyCysSerLysSer-----CysGluAlaAlaGlnTrpSer 170  
 601 ACAGCTCATTTGTTATGAACAAATGAAGATGTTTACTGTGAAGTTTCAGTATTTTGAAG 660  
 171 SerGlyPheThrValLeuGlnAlaSerileAspAlaAlaileleGlnLeuLysThrAen 190  
 661 GAAGGTTTGTGGCTCTTCAAGCTGCATTAATGCTGCTATTATAGAAATCACACAAAT 720  
 191 ValSerLeuTrpLysGluLeuGluSerThrLysAlaValileMet----- 205  
 721 CACTCAGTGATGAGGAGCTGATGTCAGTTACTTGGAAAAAATATGAAGATGCATCTCTTC 780  
 206 ---GlyGluThrAlaValValGluileAspThrPheProArgGlyValileleuileTrp 224  
 781 ATTGGTCAATCAGGAGTTATA-----ACTGATTGTACCTTTTTC 822  
 225 LeuValileAlaPheSerPropheGlyTyThrPheLeuAlaileHisileValAlaGluLys 244  
 823 TGCATTTATTTCAATTTCTTCTCATTTACTATGATCTGTTAATGTCACAGAGAGAGG 882  
 245 GluLysLysileLysGluPheLeuLysileMetGlyLeuHisAspThrAlaPheTrpLeu 264  
 883 ---AAAAGATGAAGCCTTGTATGACATGATGGTCTTGGGATTCAGCGTTCTGGCTC 939  
 265 SerTrpValLeuLeuTyThrSerLeuilePheLeuMetSerLeuLeuMetAlaValile 284  
 940 TCTGGGTTTGTCTATGCTGCTTCACTTTCATTTATGGCCCTTTCTGGCAGCTGTT 999  
 285 AlaThrAlaSerLeuLeuPheProGlnSerSerileValilePheLeuLeuPhePhe 304  
 1000 ATAAGATCTACCAGTTTATCATTTTGTCTGTGGTCTCATGTTAGTCTTACGCTCTTCTC 1059  
 305 LeuTyGlyLeuSerSerValPheAlaLeuMetLeuThrProLeuPheLysLysSer 324  
 1060 CTGTATGATATCTTTGGTAGCTTGGCTTCTTATGAGCATCTTGTAAAGAAATCT 1119  
 325 LysHisValGlyileValGluPhePheValThrValAlaPheGlyPheleGlyLeuMet 344  
 1120 TTCCTCACCGGCTGGTGTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1176  
 345 IleleleleleleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHis 364  
 1177 ACATCATGTACAGACACCTTCTGTGATCTTGGATGGATTTTAAAGTCTGCTTAGTCCC 1236  
 365 CysThrPheValileGlyileAlaGlnValMetHisLeuGlu---AspPheAsnGluGly 383  
 1237 TTTGCTTCTGCTTGGATGGCCAGCTTTTACACTTGGACTATGATTTGAATTTCTAAT 1296  
 384 AlaSerPheSerAsnLeuThrAlaGlyProTyProLeuileleThrleleMetLeu 403  
 1297 GCA---TTTCTCATCATCGGACGCTCAAAATCTCATGTAGCAACAAATTTTCATGTTG 1353  
 404 ThrLeuAsnSerilePheTyThrValLeuAlaValTrpLeuAspGlnValileProGly 423  
 1354 GCATTTGACACTTGCCTCTATCTGGCATTTGGCATTTTCTTGAATAAATTTTGGCAAT 1413  
 424 GluPheGlyLeuArgArgSerSerLeuTyThrPheLeuLysProSerTyTrpSerLysSer 443  
 1414 GAATATGACATCGACGCTCCACCTTTGTTTCTTCTGAAGCTCTATTTTGGTCTCAACA 1473  
 444 LysArgAsnTyThrGluGluLeuSerGluGlyAsnValAsnGlyAsnileSerPheSerGlu 463  
 1474 CAAAGACTGATCAGTGGGCTTGAAGATGAATGGATGCCGCTCTTCTTCTTCTATGAC 1533  
 464 IleleleGluProValSerSerGluPheValGlyGluAlaileArgleSerGlyile 483  
 1534 TCTTTTCAACAGCGCTCCAGAAATTCAGAGGAAAGAGCCATCAGAAATGTT 1593  
 484 GlnLysThrTyArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAsp 503



QY 137 TyrGluLeuArgPhe---PheProAspMetIleProValSerSerIleTyrMetAspSer 155  
Db 541 TATCATTTGAAGTTCTTGCTAGGACATGGAATGCGACAAAGAGGAGGACACAGGACCAT 600  
QY 156 ArgAlaGlyCysSerLysSer-----CysGluAlaAlaGlnTyrTrpSer 170  
Db 601 ACAGCTCATTTGTTATGAACAACATGAAGATGTTTACTGTGAAGTTTCAGTATTTTGAAG 660  
QY 171 SerGlyPheThrValLeuGlnAlaSerIleAspAlaIleIleGlnLeuLysThrAsn 190  
Db 661 GAAGGTTTTGTGGCTCTTCAAGCTGCCATTAAATGCTCTATTATAGAAATCACACAAAT 720  
QY 191 ValSerLeuTrpLysGluLeuGluSerThrLysAlaValIleMet----- 205  
Db 721 CACTCAGTATGAGGAGGCTGATGTCAGTATCTGGAATAATGAAGATGCAATCTCTTC 780  
QY 206 ---GlyGluThrAlaValAlaGluIleAspThrPheProArgGlyValIleLeuIleTyr 224  
Db 781 ATTGGTCAATCAGGAGTTATA-----ACTGATTTGTACCTTTTTC 822  
QY 225 LeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIleValAlaGluLys 244  
Db 823 TGCATTATTTCAATTTCTCATTCAATTTACTATGTCATCTGTTAATGTCACAGAGAGAGG 882  
QY 245 GluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThrAlaPheTrpLeu 264  
Db 883 ---AAAGGATGAAGGCTTGATGACCAATGATGGTCTTCGGGATTCAGCGTCTTGGCTC 939  
QY 265 SerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeuMetAlaValIle 284  
Db 940 TCCTGGGGTTTGCTCTATGCTGTTTTCATCTTCATTTATGCCCCCTTTTCITGGCACTGTT 999  
QY 285 AlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePheLeuLeuPhePhe 304  
Db 1000 ATAAGATCTACCCAGTTATCATTTTGTCTGGCTTCATGTTAGTCTTCAGCCCTTTCTC 1059  
QY 305 LeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeuPheLysLysSer 324  
Db 1060 CTGTATGATATATCTTTGGTAGCTTTGGCTTTCTTAATGAGCATCTTTGGTAAAGAAATCT 1119  
QY 325 LysHisValGlyIleValGluPhePheValThrValAlaPheGlyPheIleGlyLeuMet 344  
Db 1120 TTCCTACCGGCTGGTGGTCT 1176  
QY 345 IleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHis 364  
Db 1177 ACATCAGCTGTACAGACACCTTCCTGCATCCTTGGAGTGGATTTTAAGCTTGCCTTAGTCC 1236  
QY 365 CysThrPheValIleGlyIleAlaGluValMetHisLeuGlu---AspPheAsnGluGly 383  
Db 1237 TTTGCCCTTCATGCTTGGATGGCCAGCTTTTACCTTGGACATGATTTGAATTTCTAAT 1296  
QY 384 AlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIleIleMetLeu 403  
Db 1297 GCA---TTCTCTCATCCATCGGCGGCTCAAATCTCATTTGAGCAACAAATTTTCATGTTG 1353  
QY 404 ThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnValIleProGly 423  
Db 1354 GCATTTGACATGCTCTCTGGCATGCGATTTACTTTGAAATAATTTTGCCTCAAT 1413  
QY 424 GluPheGlyLeuAlaGlySerSerLeuTyrPheLeuLysProSerTyrTrpSerLysSer 443  
Db 1414 GAATATGACATCGACGCTCCACCTTTGTTTTCTCGAAGTCTCTCATTTTGGTCTCAACA 1473  
QY 444 LysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGlu 463  
Db 1474 CAAAAGACTGATCATCGTGGCCCTTGAAGATGAATGGAATGGAATGGAATGGAATGGAATG 1533  
QY 464 IleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIleSerGlyIle 483  
Db 1534 TCTTTTGACAGCGCTCCAGANTTCCAGGGAAGAGAGCCATCAGATCAGATCAGATGTT 1593

QY 484 GlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAsp 503  
Db 1594 ACAAAAGAAATATAAAGAAAGACCTGATAAAATGAAGCCCTTGAAGATCTGGTATTGAC 1653  
QY 504 IleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeu 523  
Db 1654 ATTTACGAAGCCAAATCCTCTGCAATCTTGGTCAAGTGGAGCTGGAAGTCAACACTG 1713  
QY 524 MetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIleTyrGlyHis 543  
Db 1714 CTAAACATTTCTAGTGGGTGCTGTTCCCAACCAAGGCTTCAGTCACCATCTATAACAAT 1773  
QY 544 ArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIleCysProGln 563  
Db 1774 AAGCTTTTCAGAAATGGCTGACCTAGAAAATCTCAGCAAGCTGACCGAGTTTGCCCAA 1833  
QY 564 LeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIle 583  
Db 1834 TCCAATGTCAATTTGACTTCTCTCTGTAAGAGAAACCTCAGACTTTTGTCTAAATA 1893  
QY 584 LysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAsp 603  
Db 1894 AAAGGATTT-----CTGCCACAGAGTGGATATA----- 1923  
QY 604 MetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeu 623  
Db 1923 ----- 1923  
QY 624 SerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAla 643  
Db 1924 -----GAGATTTCTGTTGGATGACCACTTGTCT 1953  
QY 644 GlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsn 663  
Db 1954 GGATGGATGCTCTTTTCAAGACACCAAGTATGGAACCTTCTGAAAGAAACGCAAAACAGAC 2013  
QY 664 ArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLys 683  
Db 2014 CCGGTGATCTCTCTTCTCAGTACCAGTTCATGATGAGCCGACATCTCTGGCGACAGGAA 2073  
QY 684 AlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLys 703  
Db 2074 GTATTTCTCTCCAAAGGAAGCTAAAGTGCAGCGGCTCTTCTTTGTTCTTAAAGAAAGAA 2133  
QY 704 TrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeu 723  
Db 2134 TCGGGGATGGATATCACTTAAGCTTCAGTAAATGAATATGTTGTTGAGGAAACATA 2193  
QY 724 SerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGln 743  
Db 2194 ACATCAGCTTTAAACAGACACATCCTCTGATGCCAAATTTATCAGCCAAAGCGAAGGAAA 2253  
QY 744 LeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeu 763  
Db 2254 CTATTTTATACATACCCTTAGAAGAAACAATAATTAATGAATCGGACATTTTACAGGATCTT 2313  
QY 764 AspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAsp 783  
Db 2314 GATAGCTATCTGACCTAGGAATTCAGAAATATGTTGTTTCCATGACACATTTTGAATGAA 2373  
QY 784 ValPheLeuLysLeuGluValGluAlaGluLeuAspGlnAlaAspTyrSerValPheThr 803  
Db 2374 GTATTTCTCAGACTAGAAGGAAATCTACAATTAATGAATCGGACATTTTATTTTGGGA 2433  
QY 804 GlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu 823  
Db 2434 GAAGTACAAAGCGAAAGAGCTGACGACACTGAAGGCTTGTGTAGATGGAACAGTCTCTC 2493  
QY 824 LeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGln 843  
Db 2494 TCTTCACTTACACAGATGAGAAAGACA---ATAGTGGTGTGGTCTCTCTGGCGACAGCAA 2550  
QY 844 MetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLysSerValArg 863



Db 3767 CTAGCTGCTTCATACCTCTATTTCAGACCTTTGCTTTGTTCTAAGATGCATG 3826  
Qy 1229 GluLysLysTyrGlyGlyAArgSerIleAArgLysAspProPhePheArgAsnLeuSerThr 1248  
Db 3827 GAACATAAATGTGGAAAGAAAGAAATCGGAAAGATCCTCTTTTCAGA---ATTTCCTCC 3883  
Qy 1249 LysSerLysAsnArgLysLeuProGluProAspAsnGluAspGluAspVal 1268  
Db 3884 CAAAGTAGATGCTAAG---CCAAATCCAGAGAACCCATAGATGAAGATATT 3940  
Qy 1269 LysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGluLysPro 1288  
Db 3941 CAAACAGAAAGAAATAAGACAGCAGCCATGCTGCTGACCACTTCAATCTTAGATGAGAACCT 4000  
Qy 1289 SerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSer 1308  
Db 4001 GTTATAATTGCCAGCTGCTTACACAAAGATATGCGAGCCAGAGAAAGATTGCTTTTCA 4060  
Qy 1309 ArgLysValLysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluLeu 1328  
Db 4061 AAGAGGAAGAGAAATAGCAGCAAGAAATATCTCTTCTGTGTTCAAGAAGGTGAAT 4120  
Qy 1329 LeuGluLeuLeuGlyProAsnGlyAlaGlyLysSerThrIleLeuAsnIleLeuValGly 1348  
Db 4121 TTGGATGTTAGACCCATGCTGCTGGAAGATTCATCTATTAGATGATCTGG 4180  
Qy 1349 AspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGlu 1368  
Db 4181 ATCAAAAGCAACTGCTGAGAGGTGGAATGAAAGCTGCAGTTCA----- 4228  
Qy 1369 AspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTyrProAsp 1388  
Db 4229 -----GTTTGGGCCACCTGGGTACTGGTCAAGAGACGTGCTGGCCCATG 4279  
Qy 1389 ThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAsp 1408  
Db 4280 CTGACGTTGAGGGAACACCTGGAGGTATGCTGCGCTCAAGGGCTCAGGAAGCGGAC 4339  
Qy 1409 MetLysGluValIleSerArgIleThrHisAlaLeuAspLeuGlyGluHisLeuGlnLys 1428  
Db 4340 GCGAGGCTCGCCATCGCAAGATTAAGTAGTGCTTCAAACTGTCATGACGAGCTGAATGT 4399  
Qy 1429 ThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeu 1448  
Db 4400 CCTGTGCAGAAATTAACAGCAGGAATCACGAGAAAGTTGTGTTGCTGAGCCCTCTG 4459  
Qy 1449 GlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLys 1468  
Db 4460 GGAACACTCACTGCTTGTCTGCTGATGAACCATCTACGGGCATAGACCCACAGGGCAG 4519  
Qy 1469 GlnHisMetTyrArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeu 1488  
Db 4520 CAGCAATGTGGCAGGCAATCCAGGCAGTGTGTAAACACACAGAGAGGTGCTCTCTG 4579  
Qy 1489 ThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSer 1508  
Db 4580 ACCACCCATACTGGTGGTGGGCGAAGCCTTGTGTGACCGTGTGGCCATCATGTGTCT 4639  
Qy 1509 GlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyr 1528  
Db 4640 GGAAGGCTTAGATGCTCATCTCAACACTGAAACCAACTTGAACCAACTTGGCAGGATTAC 4699  
Qy 1529 PheLeuGluLysLeuLysAspTyrIleGluAsnLeuGluValAspArgLeuGlnArg 1548  
Db 4700 ATTCAGAGCTAAAGTGAAG-----GAAACGCTCAAGTGAATTTGGTCCCACT 4750  
Qy 1549 GluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSerPheSerIleLeuAla 1568  
Db 4751 GAGATTCTGAGCTTTTCCACAGGCTGCGAGGCAAGGATTCTCTTTTGTAAACC 4810  
Qy 1569 TyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluGlu 1588

Db 4811 TATAAGCTGCCGTGGCAGACGTTTACCTCTATCAGACCTTTTCAAAATTAGAAGCA 4870  
Qy 1589 AlalysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThrLeuGluGlnVal 1608  
Db 4871 GTGAGCAATAACTTTAACTTGAAGATACAGCCTTCTCTAGTCACACTGGAGAGGTA 4930  
Qy 1609 PheValGluLeuThrLysGluGlnGluGluAspAsnSerCysGlyThrLeuAsnSer 1628  
Db 4931 TTCTTAGAGCTTCTTAAAGAACAG---GAAGTAGGAAATTTTGATGAAGAAATTTGATACA 4987  
Qy 1629 ThrLeuTyrTyrGlu 1633  
Db 4988 ACAATGAGATGAAA 5002  
RESULT 4  
PCT-US04-02188-40  
; Sequence 40, Application PC/TUS0402188  
; GENERAL INFORMATION:  
; APPLICANT: Bayer Pharmaceuticals Corporation  
; APPLICANT: Eveleigh, Deepa  
; APPLICANT: Bigwood, Douglas  
; APPLICANT: Taylor, Ian  
; TITLE OF INVENTION: EXPRESSION PROFILES FOR COLON CANCER AND METHODS OF USE  
; FILE REFERENCE: 5151  
; CURRENT APPLICATION NUMBER: PCT/US04/02188  
; CURRENT FILING DATE: 2004-01-23  
; PRIOR APPLICATION NUMBER: 60/442,582  
; PRIOR FILING DATE: 2003-01-24  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; LENGTH: 5677  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
PCT-US04-02188-40  
Alignment Scores:  
Pred. No.: 3e-295 Length: 5677  
Score: 3108.00 Matches: 672  
Percent Similarity: 60.02% Conservative: 319  
Best Local Similarity: 40.70% Mismatches: 536  
Query Match: 36.89% Indels: 124  
DB: 1 Gaps: 26  
US-10-090-458-5 (1-1642) x PCT-US04-02188-40 (1-5677)  
Qy 6 ArgGluValGlyValTyrArgGlnThrArgThrLeuLeuLysAsnTyrLeuIleLys 25  
Db 148 AGAAGATCAGTGTGTGTCACAACTTGGGCCCTTATTATGCAAGAACTTTCTTAAATAA 207  
Qy 26 CysArgThrLysLysSerSerValGlnGluLeuPheProLeuPhePheLeuPheTrp 45  
Db 208 TGGAGATGAAGAGAGAGTCTTAATGAATGGCTGAATTCATTGCTCTACTACTTGT 267  
Qy 46 LeuIleLeuLysSerMetMetHisProAsnLysLysTyrGluGluValProAsnIleGlu 65  
Db 268 TTGTATATATATCTCATGATCATCAAGTAAATGATTTTCTTCACTGCTTACCATGGAC 327  
Qy 66 LeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGlyTyrThrProValThr 85  
Db 328 CTGGACCGGTAGATACATTAAATGAATCCAGATTTTCTGTGTATACACACTGTCAAC 387  
Qy 86 AsnIleThrSerSerIleMetGlnLysValSerThrAsp----- 98  
Db 388 AACAGCCCAACAGATTAATGAATAAGTACGCTCTACTCCCTTCTGGCAGGTAAGAG 447  
Qy 99 -----HisLeuProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeu 116  
Db 448 GTCTTGGACATGCCAGATGAGAAAGTATTAAAGANTTCA----- 489  
Qy 117 ThrSerSerLeuSerLysProSerAsnPheValGlyValValPheLysAspSerMetSer 136  
Db 490 -----GCAATATTATCTCTGAAGAAATAGTAGAGTCACTTTACTATACTACTCA 540

QY 510 ThrAlaLeuLeuGlyHisSerGlyThrGlySerThrLeuMetAsnIleLeuCysGly 529  
Db 1703 ACGCAATCTGGGTACAGTGGAGCTGGCAATCTTCACTGCTAAATATCTTAATGGA 1762  
QY 530 LeuCysProProSerAspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAsp 549  
Db 1763 TTGCTGTCTCAACAGAGGATCAGTTACCATCTATAATAAAAAATCTCTCTGAATGCAA 1822  
QY 550 GluMetPheGluAlaArgLysMetIleGlyLysCysProGlnLeuAspIleHisPheAsp 569  
Db 1823 GACTTGAGGAAATCAGAAAGATAACTGGGGTCTGTCTCAATTCATGTTCATTTGAC 1882  
QY 570 ValLeuThrValGluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsn 589  
Db 1983 ATACTACCGTGAAGGAACCTCAGCTGTTGCTAAATAAAGGATTCATCTAAAG 1942  
QY 590 AsnIleLeuGlnGluValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLysAsp 609  
Db 1943 GAAGTGAACAAGAGGTACACAGGAATATTATGGAATGGACATGCAAAACATTCAGAT 2002  
QY 610 AsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeuSerGlyIleAlaVal 629  
Db 2003 AACCTTGCTAAACATTTAAGTGAAGGACAGAAAGAAAGCTGACTTTGGGATTAACAT 2062  
QY 630 LeuGlyAsnProLysIleLeuLeuLeuAspGluProThrIleLysMetAspProCysSer 649  
Db 2063 TTAGGATCTCTCAATTTTGTCTTTAGATGAACCACTACTGGATGGATCCCTTTTC 2122  
QY 650 ArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSer 669  
Db 2123 AGAGATCAAGTGTGAGCCCTCTGAGAGCGGTAGACGACATCATGTGATCCCTTTTCAGT 2182  
QY 670 ThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGly 689  
Db 2183 ACCCAGTCCATGGATGAGGTGACATCTCTGCTATAGAAAGGTGATCATGTCCCAATGGG 2242  
QY 690 MetLeuLysCysValGlySerMetPheLeuLysSerLysTyrGlyIleGlyTyrArg 709  
Db 2243 AGACTGAAGTGTGAGGTCTCTCTATSTTTTGAAGAGAGGTGGGTCTTGATATCAC 2302  
QY 710 LeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGln 729  
Db 2303 CTAAGTTTACATAGGAATGAATATGTAACCCAGAAACAAATCATCTTCACTTACAT 2362  
QY 730 HistLeuProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuPro 749  
Db 2363 CACATCCCGATGTCTAAATTAACACAGAAACAAAGAAAGCTTGTATATATCTTTGCCA 2422  
QY 750 PheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeu 769  
Db 2423 CTGGAAGGACAAATACATTTCCAGATCTTTTCACTGATCTGGATAAGTGTCTGACCAG 2482  
QY 770 GlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGlu 789  
Db 2483 GGAGTGACAGGTATGACATTTCCATTCACCTCTAAATGAAGTCTTTATGAACCTGGAA 2542  
QY 790 ValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGlu 809  
Db 2543 GCACAGTCAACTATCGAACAA---GATTTCT-----GAACAAGTGGAGATG 2584  
QY 810 GluMetAspSerLysSerPheAspGluMetClnSerLeuLeuIleLeuSerGluThr 829  
Db 2585 ATAAGAGACTCAGAAAGCCTCAATGAATGGAGCTGGCTCACTCTTCTCTCTGGAATG 2644  
QY 830 LysAlaSerLeuValSerThrMetSerLysTrpLysGlnGlnMetTyrThrIleAlaLys 849  
Db 2645 CAGACAGCT---GTGAGTGACATGGGCTCTGGAGAAATGCAAGTCTTTGGCATGGCAGCG 2701  
QY 850 PheHisPheThrLeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeu 869  
Db 2702 CTCGGTTCTTAAAGTTAAACGTCACAACTAAA-----GRTATTATGACCCCTA 2749

QY 870 LeuIlePhePheThrValGlnIlePheMetPheLeuValHis-----HisSerPhe 886  
Db 2750 TTATGGTATTGGATCGCAATATTCCCTTTGATTGTTGAAATATATATATATGCTATG 2809  
QY 887 LysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLysProGlyAsp 906  
Db 2810 TTAATGAAAGATCGATTTGGAATTTTAAACCAAGATTTTCTCTCTCTCTCTCTCTCT 2869  
QY 907 LysProHisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAspSerAspIleSer 926  
Db 2870 CTTCCCCAGGAACCCCGTACCAGCTGTTGATCATCAATAACACAGAAATCAAATTTGAA 2929  
QY 927 AspLeuIleSerPhePheThrSerGlnAsnIleMetValThrMet----- 941  
Db 2930 GATTTTATAAATCACTGAGCATCAAAATATATCTTTTGGAGTAGATGACTTTGAAAC 2989  
QY 942 IleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMetHis 961  
Db 2990 AGAATGTGATGATGGCTCTCA-----TACATGGAGTATCATAGTTTCTTGGT 3040  
QY 962 SerGluLysAspTyrValPheAlaValPheAsnSerThrMetValTyrSerLeuPro 981  
Db 3041 AAACAAAGGATTTAGATTTTTCAGTTGTGTATATACCAAGAGATTCGACTGTTTCCA 3100  
QY 982 IleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIle 1001  
Db 3101 ATTCTTATGATATATATGCAATGGGCTACTTCAATGTTTATATCACACACATATT 3160  
QY 1002 GlnIleTrpSerTrpPhePheGlnGluIleThrAspIleValPheLysIleGluLeu 1021  
Db 3161 CGAATTGAGTCAAGCCCATTT-----CCTCTAGCCACATAGGACTCTGAGCTGGT 3214  
QY 1022 TyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAlaMet 1041  
Db 3215 CCGGATGGTCTCTTTCTTTCTTTGTTCTTGTAGCTATCTCTCTTATATCACCATG 3274  
QY 1042 GluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeu 1061  
Db 3275 GGCACCATCAGTGAATTACAGAAAAATGCTAAGTCCAGCTATGGATTTCAAGGCTCTAC 3334  
QY 1062 ProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIleLeu 1081  
Db 3335 ACTCTGTTACTGTTGGTGGCAGCAGCTAGTGACGCTCAGCTTCTTCAATTTTAAATCTC 3394  
QY 1082 IleLeuMetLeuGlySerLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLys 1101  
Db 3395 CTTTAAATG-----TATTTAATTTCTACATAGAAACATGCGACTCTCTTATT 3445  
QY 1102 PheLeuAlaValValPheCysLeuIle-----GlyTyrValProSerValIle 1117  
Db 3446 ACAAGCCAAATGTTGTTGCTTGTATAGTTACTCTCTGTTATGCGAGCTTCTTGTCTC 3505  
QY 1118 LeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrp 1137  
Db 3506 TTTCTCATATATATGATATCATTTATTTTCGAAAGAGAGAAAAACAGTGGCTTGG 3565  
QY 1138 SerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePhe 1157  
Db 3565 TCAATTTACTTCTTTTGGCTCCACCATCATGTTTTCATCACTTTAATCAATCATTTT 3625  
QY 1158 ---MetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIle 1176  
Db 3626 GACCTPAGTATATGATTAACCCATG-----GTATTTGGTCTCTCA 3667  
QY 1177 TyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLys 1196  
Db 3668 TATACCTTGTCTGGA---TTTAAACCTTTTGGAAAGTG-----AGA 3706  
QY 1197 AsnValAspThrTyrAsnProTrpAspArgLeuSer----- 1208  
Db 3707 GACCAAGGAGCTACAGAGAATTTCCAGAGCAATTTTGAATGAGTGGCCACTGATTTT 3766  
QY 1209 ValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyr 1228

```

Db      4905  TTTCCCGGCGTCTCGGCGAGGAGGTTACTCTCTCTGTATGTTTATAAGTTGCCAGTG 4964
Qy      1574  GluAspValGlnSerLeuSerGlnSerPhePheLeuLeuGluGluAlaLysHisAlaPhe 1593
Db      4965  GAAGATGTGCAACCTTTAGCCCAAGCTTTCTTGAATTAGAGAAAGTTAAACAGGCGCTTT 5024
Qy      1594  AlaileGluGluTySerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThr 1613
Db      5025  GACCTAGAGAGGTACAGCCTCTCACAGTCTACCTCGGACGAGTTTCTCTGGAGCTCTCC 5084
Qy      1614  LysGluGlnGlu-----GluGluAsp 1620
Db      5085  AAGGAGCAGAGCGTGGGTGATTTTGAGGAGAT 5117

```

## RESULT 3

```

US-10-796-280-368
; Sequence 368, Application US/10796280
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001510
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 368
; LENGTH: 5296
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-796-280-368

```

## Alignment Scores:

```

Pred. No.:      6,84e-296      Length:      5296
Score:          3114.00      Matches:      687
Percent Similarity: 60.96%      Conservative: 328
Best Local Similarity: 41.26%      Mismatches: 556
Query Match:      94      Indels: 94
DB:               6      Gaps: 33

```

US-10-090-458-5 (1-1642) x US-10-796-280-368 (1-5296)

```

Qy      5  IleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLysAsnTrpLeuLeu 24
Db      182  ATGAACAGAAAGCGGTATCATGACAAACCAAGCACTCTGTGCAAGATTTCTTAAG 241
Qy      25  LysCysArgThrLysLysSerSerValGlnGluLeuLeuPheProLeuPhePheLeuPhe 44
Db      242  AAATGGAGGATGAAAGAGAGAGCTTATTGGAATGGGGCTCTCAATCTCTAGGACTG 301
Qy      45  TrpLeuLeuLeuLeuSerMetMetHisProAsnLysLysTrpGluGluValProAsnIle 64
Db      302  TGTATTGCTCTGTTTCCAGTTCCATGAGAATGTCAGTTTCTTGGAAATGGCTCTCTAG 361
Qy      65  GluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLeuGlyTrpThrProVal 84
Db      362  AAATCTGGGAAGGTAGATAAATTAATAGCTCTCTTTAATGGTTGTGTATACCAATA 421
Qy      85  ThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu-----Pro 101
Db      422  TCTAATTTAACCCAGCAGATATGATGAATAAACAGCACTGTCTCTCTTTTGAAGAAACA 481
Qy      102  AspValIleIleThrGluGluTrpThrAsnGluLysGluMetLeuThrSerSerLeuSer 121
Db      482  AGTGTCATTGGGCGCACCAATAAAACACACATGGAGCAAAATCTCTCGGAAAATTTACCA 541
Qy      122  LysProSerAsnPheValGlyValValPheLysAspSerMetSerTrpGluLeuArgPhe 141
Db      542  TATGCT-----ATGGGAATCATCTTTAATGAACATTTCTCTATAGATTAATATT 592
Qy      142  PheProAspMetIleProValSerSerIleTrpMet---AspSerArgAlaGlyCys--- 159

```

```

Db      593  TTTCCAGGATAT-----AACAGTCCACTTTGGAAGAAAGATTTCTCAGCTCATTTGCTGG 646
Qy      160  -----SerLysSerCysGluAlaAlaGlnTrpSerSerGlyPheThrVal 175
Db      647  GATGATATGGTGAAGTTTTCATGTCATGACCATGACCAATACTGGAATAGAGATTTGGCT 706
Qy      176  LeuGlnAlaSerIleAspAlaAlaIleIleGlnLysThrAsnValSerLeuTrpLys 195
Db      707  TTACAAACAGCTATTATATCTGCAATATATAAATCACAACCAATCACCTGTGATGGAG 766
Qy      196  GluLeuGluSerThrLysAlaValIleMetGlyGluThrAlaValValGluIleAspThr 215
Db      767  GAGTTGATGTCAGTTACTGCTATAACTATGAAGACATTTACCTTTCATACTAAATCTT 826
Qy      216  PheProArgGlyValIleLeuIleTrpLeuValIleAlaPheSerProPheGlyTrpPhe 235
Db      827  CTTCCACAATGAGATGTTTATTATTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 886
Qy      236  LeuAlaIleHisIleValAlaGluLysGluLysGluLysGluPheLeuLysIleMet 255
Db      887  ATATCACTCAATGTAAACAAAGAGAGA---AAAAGTCTAAGAAATTTGATGAATAATG 943
Qy      256  GlyLeuHisAspThrAlaPheTrpLeuSerTrpValLeuLeuTrpThrSerLeuIlePhe 275
Db      944  GGCTCCAGATTCAGCAATCTGGCTCTCTCTGGGCTAATCTATCTATGCTGCTCTCATCT 1003
Qy      276  LeuMetSerLeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSer 295
Db      1004  ATTATTTCCATATTCRTTACATATATCAATCAATCAATCAATCAATCAATCAATCAAT 1063
Qy      296  SerIleValIlePheLeuLeuPheLeuTrpGlyLeuSerSerValPhePheAlaLeu 315
Db      1064  TTCATGTCATATTTATCTCTTTTATATGCTTATCTTTGTTAGTCTTTGTTGTTCT 1123
Qy      316  MetLeuThrProLeuPheLysSerLysHisValGlyIleValGluPhePheValThr 335
Db      1124  CTGATGAGTGTGCTGTTAAAGAAAGCTGCTCTCACCAAATTTGTTGTTTCTCTTACC 1183
Qy      336  ValAlaPheGlyPheIleGlyLeuMetIleIleLeuLeuGluSerPheProLysSerLeu 355
Db      1184  CTCCTTTTGGGATGCTGGGA---TTCACCTGATTTTATGACCACTCTCTCTCATCTCTG 1240
Qy      356  ValTrpLeuPhe-----SerProPheCysHisCysThrPheValIleGlyIleAla 372
Db      1241  GAGTGGATTTTGAATATTTGTAGCCCTTTT-----GCCTTTACTACTGGAATGATT 1291
Qy      373  GlnValMetHisLeuGluAspPheAsn---GluGlyAlaSerPheSerAsnLeuThrAla 391
Db      1292  CAGATTTATCAACTG---GATTATACTTGAATGGTGAATTTTCTCTGACCCCTCAGGA 1348
Qy      392  GlyProTrpProLeuIleIleThrIleLeuMetLeuThrLeuAsnSerIlePheTrpVal 411
Db      1349  GACTCATATACATGATAGCAACTTTTCTATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1408
Qy      412  LeuLeuAlaValTrpLeuAspGlnValIleProGlyGluPheGlyLeuArgArg----- 429
Db      1409  CTATTGGCAATTATCTTTGACAAAATTTTACC-----TATGGAGATGAGCGCATTTAT 1462
Qy      430  SerSerLeuTrpPheLeuLysProSerTrpTrpSerLysSerLysArgAsnTrpGluGlu 449
Db      1463  TCTCTTTTATTTTCTTGAATTCATCTGTTTTCACACACCAAGAGCATTAATGCTTAAG 1522
Qy      450  LeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGluIleLeuGluProValSer 469
Db      1523  GTTATTGAGAAAGAAATCGATGCTGAGCATCCCTCTGATGATTTATTTTGAACCAAGT 1582
Qy      470  SerGluPheValGlyLysGluAlaIleArgIleSerGlyIleGlnLysThrTrpArgLys 489
Db      1583  CCTGAATTCAGAAAGAAAGAGCCATCAGAAATCAGAAATCTTAAGAGGAAATATAAGGA 1642
Qy      490  LysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAspIleTrpGluGlyGlnIle 509
Db      1643  AAATCTGAAAGTGGAGGATTTGAAGGCTTGTCTCTTTCACATATATGAAGTCAATC 1702

```

Qy	844	MctTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerIysSerValArg	863	Qy	1195	ArgLysAsnValAspThrTyrAsnProTyrAspArgLeuSerValAlaValIleSerPro	1214
Db	2832	ATTCGCGCAATTGCAAGGGTTCGCTTGTAAAGTTAAAGCATGAAAGAAAGCTCTTTTA	2891	Db	3879	CGAATGATGTACAGCCATTT-----CTGGTATTCTCTTAATTCCT	3917
Qy	864	SerValLeuLeuLeuLeuLeuLeuPhePheThrValGlnIlePheMetPheLeuValHis	883	Qy	1215	TyrLeuGlnCysValLeuLeuLeuLeuLeuTyrTyrGlnLysLysTyrGlyGly	1234
Db	2892	GCATCGTATTAATCTTAATGGCTGGATTTGGCTCTCTTCTGGAGTATACCATGGTG	2951	Db	3918	TTCTCTCATTTTATCATTTTCTTTTACTCTTCGATGCTGGAATGGAAGTTGGAAAG	3977
Qy	884	HisSerPheLysAsnAlaValProIleLeuLeuValProAspLeuTyrPheLeuLys	903	Qy	1235	ArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLys	1254
Db	2952	AAAAATATCAAAACAGTTACACC--TGGGAATCTTCTCTCATTAAGTATTCCTTGCT	3008	Db	3978	AAATCAATGAGAAAGGATCCTTTCTTTAGAAATTTCTCAAGAGTAGTAGTGTCACAA	4037
Qy	904	ProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSer	923	Qy	1255	LeuProGluProProAspAsnGluAspGluAspValLysAlaGluArgLeuLys	1274
Db	3009	CCTGGACACACCAACATGATCCCTCTCACTCACTACTGATCATCAATAAACAGGGCA	3068	Db	4038	AATCAGAGAAGCA-----GAAGGAGAGATGAAGATGTTAGATGGAAGAGTAGAGA	4091
Qy	924	AspIleSerAspLeuLeuSerPhePheThrSerGlnAsnIleMetValThrMet	941	Qy	1275	ValLysGluLeuMetGlyCysGlnCysCysGluGluLysProSerIleMetValSerAsn	1294
Db	3069	AGCATTGATGACTTTATACATCTGTGGAGCACCAGACATAGCTTTAGAGTGGATGCA	3128	Db	4092	ACAGCAATGCTTGAATTTCTACTAATTTTGATGAGAGCCAGTCATCATTTGCCAGGTGT	4151
Qy	942	-----IleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsn	958	Qy	1295	LeuHisLysGluTyrAspAspLysPheLeuLeuSerArgLysValLysLysVal	1314
Db	3129	TTTGGAACTAGAAATGGCAGATGACCCATCT-----TATATGGAGCCATCACA	3179	Db	4152	CTACGACAGGAGTATGACGGAAGAGGAGGAGGCTGTTTTTCCAAGAGAGATAAGATA	4211
Qy	959	ValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyr	978	Qy	1315	AlaThrLysTyrIleSerPheCysValLysGlyGluIleLeuLeuGlyLeuGlyPro	1334
Db	3180	GTGTGTTGTAATGAAAGAAATTACAGCTTTTCTGTAGCATGCAATGCCAAAGATTGAAT	3239	Db	4212	GCCACGAGAAATGCTCTCTCTGTAGAAAAGTGAAGTTTATAGGATATTATAGGACAC	4271
Qy	979	SerLeuProIleLeuValAsnIleSerAsnTyrTyrLeuTyrHisLeuAsnValThr	998	Qy	1335	AsnGlyAlaGlyLysSerThrIleIleAsnIleLeuValGlyAspIleGluProThrSer	1354
Db	3240	TGCTTCCAGTCTTATGGACATTTAGTAGAATGGCTACTTGAATGGTTAAACCATCA	3299	Db	4272	AATGGAGCTGTTAAAGGCACATCCATTAAAGGTATACTGGAGACACAAAACCACTGCT	4331
Qy	999	GluThrIleGlnIleTyrSerThrProPheGlnGluIleThrAspIleValPheLys	1018	Qy	1355	GlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAspAspSerLeuLys	1374
Db	3300	GTACATATCCGAATCGAAGAGTACATTTTGGGAATGACAGGACAAATCAATCGGA	3359	Db	4332	GGACAAGTGTCTACTG-----AAAGGAGCGGTGGAGGGGATGCCCTCGAG	4376
Qy	1019	IleGluLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyr	1038	Qy	1375	CysMetGlyTyrCysProGlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHis	1394
Db	3360	TTCTCGCATATATCATGTTCTGGCTG-----GTTTTAATCGAGTTCGCCACCTTAC	3413	Db	4377	TTCTTGGGGTACTGCCCTCAGAGAACCGCTGTGGCCCACTGACAGTGGAGCAGCAC	4436
Qy	1039	PheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSer	1058	Qy	1395	PheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSer	1414
Db	3414	ATTGCCATGACAGCATCGATGATTATAAGAACAGACAGCTCGGTCCAGCTACGGATTTC	3473	Db	4437	CTGGAGGTGTACCGCCGCGTGAAGGGCTGAGGAAAGGGGATGCTGAGGTTCATCACA	4496
Qy	1059	GlyLeuLeuProSerAlaTyrTyrIleGlnAlaValValAspIleProLeuPhePhe	1078	Qy	1415	ArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLeuPro	1434
Db	3474	GGACTCTCCCTCTGCTTACTGTTTGGCAGCGCTGTGGATGTTTCCCTGTACTTC	3533	Db	4497	CGTITAGTGGATCGCTCAAGCTCAGACGACAGCTGAAGTCTCCCTGGTGAACCTTGTCA	4556
Qy	1079	IleIleLeuLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyr	1098	Qy	1435	AlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThr	1454
Db	3534	TTGGTCTCGTTTTTATATATTTAATGAGCTACATTTCAACTCGAAGACGTGCTACTT	3593	Db	4557	GAGGGAATAAAGAGAAAGCTGTCTTCTCTGAGCATACTGGGGAACCCGTCAGTGGTG	4616
Qy	1099	ThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeu	1118	Qy	1455	LeuLeuAspGluProSerThrClyMetAspProLysAlaLysGlnHisMetTyrArgAla	1474
Db	3594	ACAATAATTCATATATTCAATCCCATGCTGTGTGGTTATTCTTTCCTCATCTTC	3653	Db	4617	CTTCTGGATGAGCGTCTGACCGGGATGACCCCGAGGGGAGCAGCAAAATGTGGACGCC	4676
Qy	1119	PheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTyrSer	1138	Qy	1475	IleArgThrAlaPheLysAsnArgLysArgAlaIleLeuThrThrHisTyrMetGlu	1494
Db	3654	ATGACATACGATTTCTTCTCATCTTTCGCAAGGGGAGAAAATAGTGCAATTGGTCA	3713	Db	4677	ATCCGGGCCACCTTTAGAAACACGAAAGGGGTGCCCTCTTAACACCCACTACATGCA	4736
Qy	1139	PheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMet	1158	Qy	1495	GluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIle	1514
Db	3714	TTTTGTTTCTATGTTGCTACTGATCTCTGTGGCTGATTTGGTTTCAGTATCTTTC	3770	Db	4737	GAGGCTGAGGCGGTGTGACCGATGCCATCATGGTATCTGGAGCTTGGATGTATC	4796
Qy	1159	GlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrPro	1178	Qy	1515	GlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeu	1534
Db	3771	-----GAAAGTGATATTCATTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	3818	Db	4797	GTTTCCATCCACACCTGCAAAAGCAAAATTTGGCAAGAGATTACTCTGCTGGAGATGAAGTG	4856
Qy	1179	LeuLeuGlyCysLeuIle-----SerPheIleLysIleSer-----TrpLysAsnVal	1194	Qy	1535	LysAspTyrIleGluAsnLeu-----GluValAspArgLeuGlnArgGluIleGlnTyrIle	1553
Db	3819	ATGATGGCTGTTGTTTATCTTCTCTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	3878	Db	4857	AAAG-----AACCTGGCAACAGTGGACCCCTCCATGTCAGAGATCCTGAGGCTT	4904
Qy				Qy	1554	PheProAsnAlaSerArgGlnGluSerPheSerIleLeuAlaTyrLysIleProLys	1573

Db 642 GAATTCACAGCAATATCTCGAAGAAATAGTAGAGTCACCTTTACTAATATCACTACTCA 701  
Qy 137 TyrGluLeuArgPhe---PheProAspMetIleProValSerSerIleTyrMetAspSer 155  
Db 702 TATCATTTGAAGTTCCTGCTAGGACATGGAAATGCCACCAAGAGGACGACAGGACCAT 761  
Qy 156 ArgNlaGlyCysSerLysSer-----CysGluAlaAlaGlnTyrTrpSer 170  
Db 762 ACAGCTCATTTGTTATGAACAACAAATGAAGATGTTTACTGTGAAGTTTCAGATTGTGGAG 821  
Qy 171 SerGlyPheThrValLeuGlnAlaSerIleAspAlaAlaIleIleGlnLeuLysThrAsn 190  
Db 822 GAAGTTTGTGGCTCTTCAGCTGCCATTAATGCTGCTATATAGAAATCACACAAAT 861  
Qy 191 ValSerLeuTrpLysGluLeuGluSerThrLysAlaValIleMet----- 205  
Db 882 CACTCAGTGTGAGGAGCTGATGCTACTGGAATAAATATGAGATGCAATTCCTTC 941  
Qy 206 --GlyGluThrAlaValValGluIleAspThrPheProArgGlyValIleLeuIleTyr 224  
Db 942 ATTGGTCAATCAGAGTTATA-----ACTGATTGTACCTTTTTC 983  
Qy 225 LeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIleValAlaGluLys 244  
Db 984 TGCATTATTTCATTTCTCATTTCTATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1043  
Qy 245 GluLysLysIleLysGluPheLeuLysIleMetClyLeuHisAspThrAlaPheTrpLeu 264  
Db 1044 ---AAAAGGATGAAGGCTTGATGACATGATGGTCTTCGGGATTCAGGCTTCGGCTC 1100  
Qy 265 SerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeuMetalValIle 284  
Db 1101 TCCGGGGTTGCTCTATGCTGTTTCATCTTCATATGCCCCTTTTCCTGGCACTGTT 1160  
Qy 285 AlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePheLeuLeuPhePhe 304  
Db 1161 ATAGATCTACCCAGTTATCATTTTGTCTGGCTTCATGTTAGTCTTCAGCCTCTTCTC 1220  
Qy 305 LeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeuPheLysLysSer 324  
Db 1221 CTGTATGGAATTATCTTTGGTAGCTTTGGCTTCTTAAATGAGCATCTTTGGTAAAGAAATCT 1280  
Qy 325 LysHisValGlyIleValGluPhePheValThrValAlaPheGlyPheIleGlyLeuMet 344  
Db 1281 TTCCTCACCGGCTGGTGGTCT 1337  
Qy 345 IleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHis 364  
Db 1338 ACATCACTGTACAGACACCTCTCTGTCATCTTGGAGTGGATTTAAGCTTGCTTAGTCCC 1397  
Qy 365 CysThrPheValIleGlyIleAlaGlnValMetHisLeuGlu---AspPheAsnGluGly 383  
Db 1398 TTTGCCTTCATGCTTGAATGGCCAGCTTTTACACTTGGACTATGATTTTGAATTTCTAAT 1457  
Qy 384 AlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIleIleMetLeu 403  
Db 1458 GCA---TTTCTCATTCATCGGACGGCTCAAACTCATTTAGTACACAAATTTCAATTTG 1514  
Qy 404 ThrLeuAsnSerIlePheTyrValLeuAlaValTyrLeuAspGlnValIleProGly 423  
Db 1515 GCATTTGCACCTGCTCTATCTGGCATTTGGCATTTTACTTTGAAAAAATTTTGGCAAT 1574  
Qy 424 GluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrpSerLysSer 443  
Db 1575 GAATATGACATTCAGCTCCACCTTTGTTTCTGAAAGTCCCTCATTTTGGTCTCAACA 1634  
Qy 444 LysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGlu 463  
Db 1635 CAAAAGACTGATCAGTGGCCCTTGAAGATGAAGTGGATCCGATCCCTTCATTTTCATGAC 1694  
Qy 464 IleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIleSerGlyIle 483

Db 1695 TCTTTTGAACAAGCGCTCCAGATTCCAAAGGAAAGAGCCATCAGAAATCGT 1754  
Qy 484 GlnLysThrTyrArgLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAsp 503  
Db 1755 ACAAAAGAAATATAAGGAAAGCCTGATAAAATAGAAAGCTTGAAGATCTCGTATTGAC 1814  
Qy 504 IleTyrGluGlyGlnIleThrAlaLeuGlyHisSerGlyThrGlyLysSerThrLeu 523  
Db 1815 ATTTACGAGGCAATCATCTGCAATCTTGGTACAGTGGAGCTGGAAATCAACTG 1874  
Qy 524 MetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIleTyrGlyHis 543  
Db 1875 CTAACATTTCTAGTGGTTCCTGTTCCCAAAAGGTTTCAGTCACCATATAACAAAT 1934  
Qy 544 ArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIleCysProGln 563  
Db 1935 AAGCTTTTCAGAAATGCTGCTAGAAATCTCAGCAAGCTGACCGAGTTTGTCCACA 1994  
Qy 564 LeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIle 583  
Db 1995 TCCAATGTGCAATTGGACTTCTCTCACTGTAAGAGAAACCTCAGACTCTTTGCTAAATA 2054  
Qy 584 LysGlyIleProAlaAsnAsnIleIleGlnValGlnLysValLeuLeuAspLeuAsp 603  
Db 2055 AAAGGGATTCGCCACAAGAGTGGATAAGAGATACAAAGGTTCTCTGGAATGGAA 2114  
Qy 604 MetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeu 623  
Db 2115 ATGAAAATATTCAGGATGTTCTGCTCAAACTTAAGTGGTGGACAGAAAGAGCTA 2174  
Qy 624 SerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAla 643  
Db 2175 ACCTTTGGGATGGCCATTTAGGAGATCTCAGATTTTCTGTTGGATGAACCAACTGCT 2234  
Qy 644 GlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsn 663  
Db 2235 GGACTGGATCCCTTTTCAAGACACCAAGTATGGAACTTCTGAAAGACGCAACACAC 2294  
Qy 664 ArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLys 683  
Db 2295 CGCGTGATCTCTTCAGTACCCAGTTCATGATGAGCGGACATCTCTGGCGAGAGAAA 2354  
Qy 684 AlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLys 703  
Db 2355 GTATTTCTCTCCAAAGGAGCTAAAGTGGCGGCTCTCTTTGTTTCTAAAGAGANA 2414  
Qy 704 TrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeu 723  
Db 2415 TGGGGGATGGATATCACTTAAGCTTCAGTAAATGAATATGTTGTAGGAAACATA 2474  
Qy 724 SerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsnAspGlnGln 743  
Db 2475 ACATCACTGTGTAACAGACATCCCTGATGCCAAATATATCAGCCAAAGGAGGAAAA 2534  
Qy 744 LeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeu 763  
Db 2535 CTTATTTATACATTTACCTTAGAAAGAACAAATAAATTTCCAGAACTTTTACAGGATCTT 2594  
Qy 764 AspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAsp 783  
Db 2595 GATAGCTATCTGACCTAGGAATTTAGAAATTTATGTTGTTTCCATGACAACTTTGAATGA 2654  
Qy 784 ValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThr 803  
Db 2655 GTATTTCTGAGTAGAAGAAATCTACATTTATGATCGGACATTTGCTATTTGGGA 2714  
Qy 804 GlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu 823  
Db 2715 GAAGTACAAAGGAAAAAGCTGACACATGAAAGGCTTTGTTGAGATGGAACAAAGTCTC 2774  
Qy 824 LeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGln 843  
Db 2775 TCTTCATTAACAAGATGAGAAAGACA---ATAGTGTGTGGTCTCTCTGCGACAGCA 2831

```
Db 4042 AATAAGATACCCAGAGAAATGCTCTCTGTTAGAAAAGGTGAGTTTAGGATTA 4101
Qy LeuGlyProAsnGlyAlaGlyLysSerThrIleAsnIleValGlyAspIleGlu 1351
Db 4102 TTAGGACACAATGGAGCTGTGTAAGACACATCCATTAAAGTGATACTGGAGACAAAA 4161
Qy ProThrSerGlyGlnValPheLeuGlyAspTyrSerSerClnThrSerGluAspAsp 1371
Db 4162 CCAACTCTGGACAAAGTGCTACTG-----AAAGGAGCGGTGGAGGGAT 4206
Qy SerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTyrProAspThrThrLeu 1391
Db 4207 GCCCTGGAGTCTCTGGGTACTCCCTCAGGAGACCGCTGTGGCCCAACCTGACAGTG 4266
Qy GlnClnHisPheGluIleTyrGlyAlaValGlyMetSerAlaSerAspMetLysGlu 1411
Db 4267 AGCAGACACCTGGAGGTGTACCGCGCTGAGAGGCTGAGGAAGCGGATGCTGAGTT 4326
Qy ValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLys 1431
Db 4327 GCCATCACCGTTAGTGGATGGCTCAAGCTCAGGACCCAGCTGAAGTCTCGGTGAAG 4386
Qy LysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnPro 1451
Db 4387 ACCTGTGCAGAGGAATAAAGAGAAAGCTGTCTTGTCTGAGCATCTGGGGAGCCCG 4446
Qy GlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMet 1471
Db 4447 TCAGTGTGTTCTGGATGAGCGCTGACCGGATGAGCCCGGAGGAGCAGCAATG 4506
Qy TrpArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaIleLeuThrHis 1491
Db 4507 TGGCAGGCCATCCGGCCACCTTTAGAAACACGGAAGGGGTGCTCTCTAACCCACCAC 4566
Qy TyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeu 1511
Db 4567 TACATGGCAGAGGTGAGGCGGTGTGACCGAGTGGCCATCATGTATCTGGAGGTG 4626
Qy ArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGlu 1531
Db 4627 AGATGTATCGTTCATCCACACCTGAAAGCAAAATTTGGCAAGATTACCTGCTGGAG 4686
Qy IleLysLeuLysAspTrpIleGluAsnLeu--GluValAspArgLeuGlnArgGluIle 1550
Db 4687 ATGAAGGTGAAG-----AACCTGGCAAGTGGAGCCCTCCATGCGAGATC 4734
Qy GlnTyrIlePheProAsnAlaSerArgGlnGlnSerPheSerSerIleLeuAlaTyrLys 1570
Db 4735 CTGAGGCTTTTCCCCAGGCTGCTCGGCAAGGAGTACTCTCTGTGTTTATAAG 4794
Qy IleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluAlaLys 1590
Db 4795 TTGCGAGTGAAGATGTGCAACCTTTAGCCCAAGCTTTCTTCAATTAGAAGAGTTAA 4854
Qy HisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheVal 1610
Db 4855 CAGAGCTTTGACCTTAGAGGAGTACAGCTCTCACAGCTACCTCGGAGGAGTTTCTCTG 4914
Qy GlnLeuThrLysGluGlnGlu-----GluGluAsp 1620
Db 4915 GAGCTCTCAATGACGAGGAGTGGGTGATTTTGGAGGAGAT 4956
```

## RESULT 2

```
PCT-US04-07967-57
; Sequence 57, Application PC/TUS0407967
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;
; APPLICANT: LEE, Ernestine A.; RICHARDSON, Thomas W.;
; APPLICANT: ELLIOTT, Vicki S.; GRIFFIN, Jennifer A.;
; APPLICANT: HAFALIA, April J.A.; DING, Li;
; APPLICANT: CHAWLA, Narinder K.; YANG, Junming;
; APPLICANT: BAUGHN, Mariah R.; THORNTON, Michael B.;
```

```
; APPLICANT: GIETZEN, Kimberly J.; MARQUIS, Joseph P.
; APPLICANT: TRIBOULEY, Catherine M.; ISON, Craig H.;
; APPLICANT: RAMKUMAR, Jayalaxmi; JIN, Pei;
; APPLICANT: CHIEN, David; BECHA, Shanya D.;
; APPLICANT: WILSON, Amy D.; HAWKINS, Phillip R.;
; APPLICANT: KHARE, Reena; BHATIA, UMESH G.
; APPLICANT: BLAKE, Julie J.; BURRILL, John D.;
; APPLICANT: HO, Anne; LEE, Sally;
; APPLICANT: ZHENG, Wenjin;
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PF-1517 PCT
; CURRENT APPLICATION NUMBER: PCT/US04/07967
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: US 60/456,169
; PRIOR FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: US 60/462,271
; PRIOR FILING DATE: 2003-04-10
; PRIOR APPLICATION NUMBER: US 60/465,412
; PRIOR FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: US 60/475,073
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/476,492
; PRIOR FILING DATE: 2003-06-06
; PRIOR APPLICATION NUMBER: US 60/502,391
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: US 60/524,102
; PRIOR FILING DATE: 2003-11-21
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PERL Program
; SEQ ID NO 57
; LENGTH: 5396
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7520474CB1
PCT-US04-07967-57
```

## Alignment Scores:

```
Pred. No.: 3,19e-307 Length: 5396
Score: 3229.00 Matches: 689
Percent Similarity: 61.60% Conservative: 328
Best Local Similarity: 41.73% Mismatches: 550
Query Match: 38.32% Indels: 84
DB: 1 Gaps: 24
```

US-10-090-458-5 (1-1642) x PCT-US04-07967-57 (1-5396)

```
Qy 6 ArgGluValGlyValT-pArgGlnThrArgThrLeuLeuLysAsnTyrLeuIleLys 25
Db 309 AGAAGATCATGCTGTGTCTACAAACCTGGGCTTATTATGCAAGAACTTCTTAAAAA 368
Qy 26 CysArgThrLysLysSerSerValGlnGluLeuPheProLeuPhePheTrp 45
Db 369 TGGAGATGAAGAGAGAGTCTCTTAATGAATGGCTGAATTCATTGCTCTACTTGT 428
Qy 46 LeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGluValProAsnIleGlu 65
Db 429 TTGTATATATATCTCATAGTCATCAAGTAATGATTTTCTCAGTCCTTACCATGAC 488
Qy 66 LeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGlyTyrThrProValThr 85
Db 489 CTGGAGCGGTAGATACATATTTAATGAATCCAGATTTTCTGTGTATACACACCTGTCC 548
Qy 86 AsnIleThrSerSerIleMetGlnLysValSerThrAsp----- 98
Db 549 AACACGACCCACAGATATATGAATAAGTAGCTCTACTCCCTCTCTGGCAGGTAAGAG 608
Qy 99 -----HisLeuProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeu 116
Db 609 GTCTGGGACTGCCAGAT-----GAGGAAGATATTACA 641
Qy 117 ThrSerSerLeuSerLysProSerAsnPheValGlyValValPheLysAspSerMetSer 136
```

Qy	604	MetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLysArgLysLeu	623
Db	1954	ATGAATAATATTCAGGATGTTCTTGCTCAAAACTTAAGTGGTGCACAGAAAAGAGCTA	2013
Qy	624	SerLeuGlyIleAlaValLeuGluYasnProLysIleLeuLeuLeuAspGluProThrAla	643
Db	2014	ACCTTTGGGATGGCATTTTAGAGATCCTCAGATTTTCTTGTTGGATGAACCAACTGCT	2073
Qy	644	GlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsn	663
Db	2074	GGATGGATCCCTTTTCAAGACACCAAGATAGGAACCTTCTGAAAGAACGCAAAACAGCA	2133
Qy	664	ArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLys	683
Db	2134	CGCTGTATCCTCTTCAGTTACCGAGTTCAATGATGAGGCCGACATCTCTGGCGGACAGAAA	2193
Qy	684	AlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLys	703
Db	2194	GTATTTCTCTCCCAAGGAAAGCTAAAGTCGCGCGGCTCTCTTTTGTCTTCAAGAGAAA	2253
Qy	704	TrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeu	723
Db	2254	TGGGGGATGGATATCATCTTAAGCTTGCAGTTAATGAATATGTGTGGAGAAAACATA	2313
Qy	724	SerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGln	743
Db	2314	ACATCACTTGTATAACACGACGCTCCCTGATGCCAAATTTATCAGCCAAAAGCAAGGAAA	2373
Qy	744	LeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeu	763
Db	2374	CTTATTTATCATTTACCTTTAGAAGAACAAATAAATTTTCAGAACCTTTACAGGATCTT	2433
Qy	764	AspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAsp	783
Db	2434	GATAGCTATCTGCACCTAGGAATTGAGAAATATGCTGTTCCATGACAACTTTCAATGAA	2493
Qy	784	ValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThr	803
Db	2494	GTATTCCTGAGCTAGAGGAAATATCAATTAATGAATCGGACATGCTATTATTGGGA	2553
Qy	804	GlnGlnProLeuGluGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu	823
Db	2554	GAAGTACAGGGGAAAAAGCTGACGACACTGAAAGGCTTGTTGTGATGGAAACAGTCCTC	2613
Qy	824	LeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGln	843
Db	2614	TCTTCACCTTACAAGATAGAGAAAGACA--ATAGGTGGTGGCTCTCTGGCGACAGCAA	2670
Qy	844	MetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLysSerValArg	863
Db	2671	ATCTGCGCAATTGCAAGGGTTCGCTTGTTAAAGTTAAAGCATGAAGAAAGCTCTTTTA	2730
Qy	864	SerValLeuLeuLeuLeuLeuLephePheThrValGlnIlePheMetPheLeuValHis	883
Db	2731	GCATCTGTTTAATTTCAATGGCTGGATTTGGCCTCTTCTTGGAGTATACCATGGTG	2790
Qy	884	HisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLys	903
Db	2791	AAATATATCAAAACAGTTACACC--TGGGAACCTTCTCCTCATTTGTATTTCTTGCT	2847
Qy	904	ProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSer	923
Db	2848	CCTCGACAACAACACATGACCCCTCTCACTCACTACTGATCATCAATAAAACAGGGGCA	2907
Qy	924	AspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThrMet----	941
Db	2908	AGCATTTGATGCTTTATACAGTCTGTGGACACCAGACATAGCTTTAGAAGTGGATGCA	2967
Qy	942	-----IleAsnAspSerAspTyrValSerValAlaProHisSerAla	955
Db	2968	TTTGGAACTAGAAATGGCACAGGTGACCCATCTTAT-----AATGGA	3009

956 AlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaValPheAsnSerThr 975  
D5 : : : : :  
3010 GCCATCAGCTGGTTGTAAATGAAGAATAACAGCTTTTCGTAGCATGCATGCCAAA 3069  
QY : : : : :  
976 MetValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeu 995  
D5 : : : : :  
3070 AGATTGAATTCCTCCAGTCTCTATGGCACTGTTAGTAATCGGCTACTTCGAATGGTT 3129  
QY : : : : :  
996 AsnValThrGluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIle 1015  
D5 : : : : :  
3130 AAACCAATCAGTACATATCCGAATCGAAGAAGATACATTTTTTGAGAGAAATGCAGGACAAT 3189  
QY ValPheIlyIleGluLeuTyrPheGlnAlaLeuGluIlyIleIleValThrAlaMet 1035  
D5 : : : : :  
3190 CCAATCGGATTCCTGGCATATATCATGTTCTGGCTG-----GTTTAACTCGAGTTGC 3243  
QY ProProTyrPheAlaMetGluAsnAlaGluAsnHisIlyIleLysAlaTyrThrGlnLeu 1055  
D5 CCACCTTACATTGCCATTGAGCAGCATCGATGATTAATAAGAACAGAGCTCGTCCCAGCTA 3303  
QY LysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIlePro 1075  
D5 : : : : :  
3304 CGGATTTCCGGACTCTCCCTCTCTGTTACTGTGTTGGCAGCGCTGTGGATGTTTCC 3363  
QY LeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisIlyrGlyLeu 1095  
D5 CTGTACTTCTGTGCTCGTTTTATATATTAATGAGCTACATTTCAAACITCGAAGAC 3423  
QY TyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSer 1115  
D5 ATGCTACTTACAATAATTCATATATTCCAATCCCATGCTGTGTGGTATTTCCTTTTCC 3483  
QY ValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGlu 1135  
D5 : : : : :  
3484 CTACATCTCATGATACAGTAGATTCTCTCATCTTTCGCAAGGGGAGAAAAATAGTGGC 3543  
QY PheTrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThr 1155  
D5 ATTGTGGTCATTTGGTTTATGTTGTCACTGTATTCTCTGTGGCTGGATTGGGTTTCAGT 3603  
QY PhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIlePro 1175  
D5 ATCTTC-----GAAAGTGATATCCATTATCTCACTTCTTCTATATATACCA 3648  
QY IleTyrProLeuLeuGlyCysLeuIle-----SerPheIleLysIleSer-----Trp 1191  
D5 : : : : :  
3649 CCTGCCACATGATTGGCTGTTGTTCTTACCTTCTCACTTCTCTTCTTCTCTCTCTT 3708  
QY LysAsnValArgLysAsnValIleAspThrTyrAsnProTrpAspArgLeuSerValAlaVal 1211  
D5 : : : : :  
3709 TCTGAAGAACGAATGATGTACAGCCATTT-----CTGTGATTC 3747  
QY IleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysLys 1231  
D5 : : : : :  
3748 CTAAATCCCTTCTCTCATTTATCATTTTTTTTCTTTTACTCTTCGATGTGGAAATGGAA 3807  
QY TyrGlyGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLys 1251  
D5 : : : : :  
3808 TTTGGAAAGAAATCAATGAGAAAGGATCCTTCTTTTGAATTTCTCCAAAGAGTAGTAGAT 3867  
QY AsnArgLysLeuProGluProProAspAsnGluAspGluAspValLysAlaGlu 1271  
D5 : : : : :  
3869 GTGTGTCAAAATCCAGAAAGACCA-----GAAGGAGAGGATGAAGATGTTCAGATGGAA 3921  
QY ArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGluLysProSerIleMet 1291  
D5 : : : : :  
3922 AGAGTGAGAACACAAATGCCCTTGAATTTCTACTAATTTTGGTGAGAGCCAGTCATCAT 3981  
QY ValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSerArgLysVal 1311  
D5 : : : : :  
3982 GCCAGTGTCTACGCAAGAGATATGCAGGAACAGGAAAGGCGTGTGTTTTTCCAAAGAGGAA 4041  
QY LysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIleLeuGlyLeu 1331



Db	823	TGCATTATTTCA	TTTTCCCTCAT	TTTCATTTCAT	TACTATG	CATCTG	CTTAA	TGTCA	CAAGACAGG	888
Qy	245	Glulysylsilelys	GluPheLeuLys	IleMetClyLeuHis	AspThrAla	PheTrpLeu	264			
Db	883	---AAAAGTGAT	GAAGCGCTT	GTATGATGGT	CTTCGG	AATTCAGG	CTTCTG	CGCT	939	
Qy	265	SerTrpValLeuLeu	ValTyrThrSerLeu	IlePheLeuMet	SerLeuLeuMet	AlaValIle	284			
Db	940	TCTTGGGGTTG	TCTATGCTGGT	TTCATCTT	CATTATG	CGCCCTT	TCTT	GGCAC	TTGTT	999
Qy	285	AlaThrAlaserLeu	PhePheProGlnSer	SerSeriValIle	ValIlePheLeu	PheLeuPhePhe	304			

[illegible]

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

1177	ACATCACTGACAGACACACCTCTCTGCGATCTCTGGAGTGGATTTTAAAGCTTGCTTAGTCCC	1236
365	CysThrPheValIleGlyIleAlaGlnValMetHisLeuGlu---AspPheAsnGluGly	383
1237	TTTGCCCTCATGCTTGGAAATGCCAGCTTTACACITGGACTATGATTGCAATCTCAAT	1296
384	AlaSerPheSerAsnLeuThrAlaGlyProTyrPProLeuIleIleThrIleIleMetLeu	403
1297	GCA---TTTCCTCATCATCGGACGGCTCAAAATCTCATTTAGCAACAAATTTTCATGTG	1353
404	ThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnValIleProGly	423
1354	GCATTTGACACTTGCCCTCTATCTGGCAATGGCGATTTACTTTGAAAAAATTTTCCCAAT	1413
424	GluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrpSerLysSer	443
1414	GAATATGACATCGACGCTCCACCTTTGTTTTCTGAACTCCTCATTTTGGTCTCAACA	1473
444	LysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGlu	463
1474	CAAAAGACTGATCACGTGGCGCTTGAAGATGAATGATGCGGATCTTCATTTTCATGC	1533
464	IleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIleSerGlyIle	483
1534	TCITTTGAAACAGCGCCTCCAGAATTTCCAAGGGAAAGAGCCATCAGAAATGTT	1593
484	GlnLysThrTyrArgLysLysGlyGlyAsnValGluAlaLeuArgAsnLeuSerPheAsp	503
1594	ACAAAAGAATATAAAGAGAAAGCCCTGTAAATATAGAAGCCTTGAAGATCTGGTATTG	1653
504	IleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeu	523
1654	ATTTCAGAGGCCAAATCATCTGCATACTTGGTCACAGTGGAGCTGGAAAGTCAACACTG	1713
524	MetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIleTyrGlyHis	543
1714	CTAAACATTTCTTAGTGGGTTGTCTGTCTCCACCACCAAGGTTTCAGTCACCATCTATAA	1773
544	ArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIleCysProGln	563
1774	AAGCTTTCAAAATGGCTGACCTACCTAGAAAAATCTCAGCAAGCTGACCGAGTTTGTCCA	1833
564	LeuAspIleHisPheAspValLeuThrValIleGluAsnLeuSerIleLeuAlaSerIle	583
1834	TCCAAATGTGAATTTGACTTCTCTCTCTGTAAAGAAAAACCTCAGACTCTTGCTAAATA	1893
584	LysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeuAspLeuAsp	603
1894	AAAGGGATTCTGCCACAAGAAGTGGATAAAGAGATACAAAGGGTTCTGCTGGAATTTGGA	1953

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 12, 2004, 21:52:05 ; Search time 325 Seconds  
(without alignments)  
3300.409 Million cell updates/sec

Title: US-10-090-458-5  
Perfect score: 8426  
Sequence: 1 MSTAIREVGWVRQRTLLK.....CGTLNLTWERTQEDRVVF 1642

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1355633 seqs, 326623912 residues

Total number of hits satisfying chosen parameters: 2711266

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-Q=/cgn2\_1/USP01.epool/US10090458/runat\_08042004\_121554\_25634/app\_query.fasta\_1.1799  
-DB=Pending Patents NA New -QFMT=fastap -SUFFIX=irpn -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10090458 -CGN1\_1\_75 @runat\_08042004\_121554\_25634 -NCPU=6 -ICPU=3  
-NO MWAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSFLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA New:  
1: /cgn2\_6/ptodata/2/pna/PCT NEW COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06 NEW COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07 NEW COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US08 NEW COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US09 NEW COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US10 NEW COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US60 NEW COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3231	38.3	5235	1	PCT-US04-07967-55
2	3229	38.3	5396	1	PCT-US04-07967-57
3	3114	37.0	5296	6	US-10-796-280-368
4	3108	36.9	5677	1	PCT-US04-02188-40
5	3108	36.9	5677	1	PCT-US04-07268-45
6	3108	36.9	5677	6	US-10-754-425-40
7	3108	36.9	5677	6	US-10-788-792-45
8	2967.5	35.2	5185	6	US-10-796-280-369
9	1475	17.5	9497	6	US-10-465-498-91
10	1473.5	17.5	9593	6	US-10-465-498-94
11	1458.5	17.3	9495	6	US-10-465-498-96

12	1439	17.1	10026	6	US-10-796-307-75
13	1439	17.1	10247	6	US-10-796-307-77
14	1439	17.1	10405	6	US-10-796-307-78
15	1363	16.2	6705	5	US-09-032-438C-5
16	1352	16.0	6705	6	US-10-336-215A-5
17	1352	16.0	6705	6	US-10-336-219A-5
18	1350	16.0	6705	6	US-10-340-097B-5
19	1350	16.0	6819	5	US-09-032-438C-2
20	1350	16.0	7488	5	US-09-032-438C-1
21	1339	15.9	6701	6	US-10-603-187-1
22	1337	15.9	6819	6	US-10-336-215A-2
23	1337	15.9	6819	6	US-10-336-219A-2
24	1337	15.9	6819	6	US-10-340-097B-2
25	1337	15.9	7318	6	US-10-783-528-12
26	1337	15.9	7783	6	US-10-336-215A-1
27	1337	15.9	7783	6	US-10-336-219A-1
28	1337	15.9	7783	6	US-10-340-097B-1
29	1295.5	15.4	6027	6	US-10-775-920-5
30	1295.5	15.4	6174	6	US-10-775-920-6
31	1295.5	15.4	6588	6	US-10-775-920-4
32	1295.5	15.4	6704	6	US-10-775-920-3
33	1291.5	15.3	6704	6	US-10-775-920-1
34	1249.5	14.8	5811	6	US-10-100-683-3889
35	546	6.5	1879	6	US-10-775-920-7
36	479	5.7	1501	6	US-10-775-920-2
37	407	4.8	4254	7	US-60-554-981-1303
38	400.5	4.8	4646	6	US-10-775-169-198
39	399	4.7	4192	6	US-10-782-413-53
40	395.5	4.7	4669	6	US-10-680-516-1
41	365	4.3	6129	6	US-10-769-619-1
42	365	4.3	6129	6	US-10-782-413-44
43	364	4.3	4572	6	US-10-790-273-1
44	360.5	4.3	6126	6	US-10-769-619-3
45	358	4.2	6129	6	US-10-769-619-5

ALIGNMENTS

RESULT 1

PCT-US04-07967-55  
Sequence 55, Application PC/TUS0407967  
GENERAL INFORMATION:  
APPLICANT: INCYTE CORPORATION; TANG, Y. Tom;  
APPLICANT: LEE, Ernestine A.; RICHARDSON, Thomas W.;  
APPLICANT: ELIOTT, Vicki S.; GRIFFIN, Jennifer A.;  
APPLICANT: HAPALIA, April J.A.; DING Li;  
APPLICANT: CHAWLA, Narinder K.; YANG, Junming;  
APPLICANT: BAUGHN, Mariah R.; THORNTON, Michael B.;  
APPLICANT: GIETZEN, Kimberly J.; MARQUIS, Joseph P.;  
APPLICANT: TRIBOULEY, Catherine M.; ISON, Craig H.;  
APPLICANT: RAKUMAR, Jayalaxmi; JIN, Pei;  
APPLICANT: CHIEN, David; BECHA, Shanya D.;  
APPLICANT: WILSON, Amy D.; HAWKINS, Phillip R.;  
APPLICANT: KHARE, Reena; BHATIA, UMESH G.  
APPLICANT: BLAKE, Julie J.; BURRILL, John D.;  
APPLICANT: ZHENG, Wenjin  
TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS  
FILE REFERENCE: PF-1517 PCT  
CURRENT APPLICATION NUMBER: PCT/US04/07967  
PRIOR FILING DATE: 2003-03-14  
PRIOR APPLICATION NUMBER: US 60/456,169  
PRIOR FILING DATE: 2003-03-16  
PRIOR APPLICATION NUMBER: US 60/452,271  
PRIOR FILING DATE: 2003-04-10  
PRIOR APPLICATION NUMBER: US 60/465,412  
PRIOR FILING DATE: 2003-04-25  
PRIOR APPLICATION NUMBER: US 60/475,073  
PRIOR FILING DATE: 2003-05-30  
PRIOR APPLICATION NUMBER: US 60/476,492  
PRIOR FILING DATE: 2003-06-06  
PRIOR APPLICATION NUMBER: US 60/502,391  
PRIOR FILING DATE: 2003-09-12

QY	1238	glyaspProPheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluPr	1258
Db	1711	AAAAGATCCCTTTTTCAGAAACCTTTCAACGAAGTCTAAAATAGGAGCTTCAGAAC	1770
QY	1258	oProAspAsnGluAspGluAspGluAspValiAlaGluArgLeuLysValLysGluLe	1278
Db	1771	ACCAGACATGAGGATGAGATGAGATGATCAAGCTCAAGAGCTCAAGAGCT	1830
QY	1278	uMetGlyCysGlnCysCysGluGluLysProSerLleMetValSerAsnLeuHisLysGl	1298
Db	1831	GATGGTTGCCAGTCTTGTGAGGAGAACCATCCATTATGCTCAGCAATTTGCATAAGA	1890
QY	1298	uTyraAspLysLysAspPheLeuSerArgLysValLysValAlaThrLysTy	1318
Db	1891	ATATGATGACAGAAAGATTTCTTTCTTTCAAGAAAGTAAAGAGTGGCACTAATA	1950
QY	1318	rLysSerPheCysValLysLysGlyGluLleLeuGlyLeuLeuGlyProAsnGlyAlaGl	1338
Db	1951	CATCTCTTTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGGTCCAAATGGTGTGG	2010
QY	1338	lysserThrIleLleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPh	1358
Db	2011	CAAAAGCACATTTATATATTTCTGGTTGGTGATATTGAACCACTTCAGGCCAGGTATT	2070
QY	1358	eLeuGlyAspTyrsrSerGluThrSerGluAspAspSerLeuLysCysMetGlyTy	1378
Db	2071	TTTAGGAGATTATTTCTTCAGAGACAAGTGAAGTGAATGATGATTTCACTGAAGTGTATGGGTTA	2130
QY	1378	rCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTy	1398
Db	2131	CTGCTCTCAGATAAACCTTTTGTGCCAGATACATTCAGGAGACATTTTGAATTTA	2190
QY	1398	rcLyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHi	1418
Db	2191	TGGAGCTGTCAAAGGAATGAGTGAAGTGAATGAAGAGTCAATAGTGAATGAACACA	2250
QY	1418	sAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLy	1438
Db	2251	TGCACCTTGATTTAAAGAACATCTTTCAGAGACTGTAAAGAACTTACCTGCAGGAATCAA	2310
QY	1438	sArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGl	1458
Db	2311	ACGAAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGA	2370
QY	1458	uProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAl	1478
Db	2371	ACCATCTACAGGTATGGATGCCAAAGCCAAACAGCACATGTGGCGAGCAATTCGAACCTGC	2430
QY	1478	aPheLysAsnArgLysArgAlaAlaIleLeuThrThrHisTyrrMetGluGluAlaGluAl	1498
Db	2431	ATTTAAACACAGAAAGCGGGCTGCTATTCTGACCACTCATATATGGAGAGGAGAGGC	2490
QY	1498	aValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGl	1518
Db	2491	TGCTGTGATCGAGTAGCTATCATGGTGTCTGGCGAGTTAAGATGATTCGGACAGTACA	2550
QY	1518	nhisLeuLysSerLysPheGlyLysGlyTyrrPheLeuGluIleLysLeuLysAspTrpIl	1538
Db	2551	ACATCTTAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTTAAATTTGAAGGACTGGAT	2610
QY	1538	eGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrrIlePheProAsnAlaSe	1558
Db	2611	AGAAAACTTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCCAAG	2670
QY	1558	rArgGlnGluSerPheSerSerLleLeuAlaTyrrLysIleProLysGluAspValGlnSe	1578
Db	2671	CCGTGAGGAAAGTTTCTTCTATTTTTGGCTCATAAATTCCTAAGGAAGATGTCAGTC	2730
QY	1578	rLeuSerGlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTy	1598
Db	2731	CCTTTTCAACAATCTTTTAAAGTGGAGAGAGCTAAACATGCTTTTGGCATTTGAAGAATA	2790
QY	1598	rSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGl	1618

Search completed: April 13, 2004, 04:14:16  
Job time : 7815 secs

Db	2791	TAGCTTTTCTCAAGCAACATTGGAACAGGTTTTTTAGAACTCCTTAAAGAACAAAGAGA	2850
QY	1618	uGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAs	1638
Db	2851	GGAAGATAATAGTTGTGGAACTTTAAACAGCACACTTTGGTGGAAACGAACACAAGAAGA	2910
QY	1638	parGValValPhe	1642
Db	2911	TAGAGTAGTATTT	2923



QY 920 rAla-----AspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMe 938  
Db 2491 TGCTGGTGAGAACTCAGATATCAGTGATCTATTAGCTCTTTTCAAGACCCAGACATAAT 2550  
QY 938 tValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAs 958  
Db 2551 G----- 2551  
QY 958 nValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTy 978  
Db 2552 -----GACTATGTTTTGTCAGCTGTTTTCAACAGTACTATGGTTTA 2592  
QY 978 sSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValTh 998  
Db 2593 TTTCTTACCTATATTAGTGAATATCATTTAGTAACCTACTATCTTTATCATTTAAATGTGAC 2652  
QY 998 rGluThrIleGlnIleTyrSerThrProPhePheGlnGluIleThrAspIleValPheLy 1018  
Db 2653 TGAACACCATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTAAGTATAGTTTTTAA 2712  
QY 1018 stleGluLeuTyrPheGlnAlaAlaLeuGluIleIleValThrAlaMetProProTy 1038  
Db 2713 AATTGAGCTGATTTTCAAGCAGCTTTGCTTGGATCATTTGTTACTGCAATGCCACCTTA 2772  
QY 1038 rPheAlaMetGluAsnAlaGluAsnHisIleLysAlaTyrThrGlnLeuLysLeuSe 1058  
Db 2773 CTTTGGCATGGAAAAATGCAGAGATCATTAAGATCAAGCTTATATCAACTTAACTTTC 2832  
QY 1058 rGlyLeuLeuProSerAlaTyrTyrIleGlyGlnAlaValValAspIleProLeuPhePh 1078  
Db 2833 AGGCTTTTGCCATCTGCATATTGGATTGGACAGCTGTTGTTGATATCCCTTATTTT 2892  
QY 1078 etleleLeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTy 1098  
Db 2893 TATCATCTTATTTGATCTAGGAAGCTTATTGGCAITTTCAATTATGATTATATTTTA 2952  
QY 1098 rThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLe 1118  
Db 2953 TACTGTAAGTTCCTTCTGCTGGTTTTTGGCTTATTTGTTGTTGTTGTTGTTGTTGTTGTT 3012  
QY 1118 uPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTyrSe 1138  
Db 3013 GTTCACCTTATATTGCTTCTTTTCACTTTTAAAGAAATTTTAAATATCAAGAAATTTTGGTC 3072  
QY 1138 rPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMe 1158  
Db 3073 ATTTATCTATTCTGTGGCAGCGTTGGCTTGTATTGCAATCACTGAAATAAATCTTTCTTAT 3132  
QY 1158 tGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIlePheIleTyrPr 1178  
Db 3133 GGGATACACAATTGCAACTATTCTTCAATATGCTTTTGTATCATCATTCCAATCTATCC 3192  
QY 1178 oLeuLeuGlyCysLeuIleSerPheIleLysIleSerTyrLysAsnValArgLysAsnVa 1198  
Db 3193 ACTTCTAGGTGCTGATTTCTTTTCAATAAGATTTCTTGGAGAATGTACGAAAAATGT 3252  
QY 1198 lAspThrTyrAsnProTyrAspArgLeuSerValAlaValIleSerProTyrLeuGlnCy 1218  
Db 3253 GGACACCTATATCCATGGAGTAGCTTTCAGTAGCTGTTATATGCTCTTACCTTGCAGTG 3312  
QY 1218 sValLeuTyrIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyLysArgSerIleAr 1238  
Db 3313 TGTACTGTGGATTTTCTCTTCAATACTATGAGAAAAAATATGGAGGCAGATCAATAAG 3372  
QY 1238 gLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluPr 1258  
Db 3373 AAAAGATCCCTTTTTCAG-----TCTAAAAATAGGAAGCTTCNGAACC 3416  
QY 1258 oProAspAsnGluAspGluAspValLysAlaGluArgLeuLysValLysGluLe 1278  
Db 3417 ACCAGACATAGGATGAAGATGAAGATGCTCAAGCTCAAGACTTAAAGGTCAAAGAGCT 3476

QY 1278 uMetGlyCysGlnCysGluLysProSerIleMetValSerAsnLeuHisLysG1 1298  
Db 3477 GATGGGTTGCCAGTGTTGTGAGGAGAAACCATTCATTATATGGTCAGCAATTTGCATAAGA 3536  
QY 1298 uTyrAspAspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTy 1318  
Db 3537 ATATGATGACAAGAAAGATTTTCTTTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAATA 3596  
QY 1318 rIleSerPheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaG1 1338  
Db 3597 CATCTCTTCTGTGTGAAAAAGGCT--ATCTAGGACTATTGGGTCCAAATGGTCTGG 3654  
QY 1338 YLysSerThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPh 1358  
Db 3655 CAAAAGCACAAATTAATATCTGTTGGTGTATTTGAACCAACTTCAGGCCAGGTATT 3714  
QY 1358 eLeuGlyAspTyrSerGluThrSerGluAspAspSerLeuLysCysMetGlyTy 1378  
Db 3715 TTTAGAGATTATCTTCAGACAGTGAAGTGAATGATTCACCTGAGTGTATGGTTA 3774  
QY 1378 rCysProGlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTy 1398  
Db 3775 CTGCTCTCAGATAAACCTTTGTGGCCAGATCTACATTGCAGAAACATTTTGAATTTA 3834  
QY 1398 rGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHi 1418  
Db 3835 TGGAGCTGTCAAAGGAATGAGTGCAGATGACATGAAGAAGTCAATAGTCGAATAACA 3894  
QY 1418 sAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleTy 1438  
Db 3895 TGCACITGATTTAAAGACATCTTCAGAAGACTGTAAAGAAACTACCTGCAGGAATCAA 3954  
QY 1438 sArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspG1 1458  
Db 3955 ACAGAAAGTTGTTGTTTGTCTTAAGTATGCTAGGAATCTCAGATTACTTTGCTAGATGA 4014  
QY 1458 uProSerThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAla-IleArgThrA 1478  
Db 4015 ACCATCTACAGGTATGGATCCCAAGCCAAACAGCACATG-----TGCAATTCGAACGTG 4068  
QY 1478 laPheLysAsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluA 1498  
Db 4069 CATTTAAAAACAGAAACGGGCTGCTTATTTCTGACCCTCCTCATATATGGAGGAGGAGG 4128  
QY 1498 laValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValG 1518  
Db 4129 CTGCTGTGATCGAGTAGCTATCATGCTGCTGGCAGTTAGATGTAACGACAGTAC 4188  
QY 1518 lHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTyrI 1538  
Db 4189 AACATCTAAAGAGTAAATTTGGAAGGCTACTTTTTTGGAAATTAATTTGAAGGACTGGA 4248  
QY 1538 leGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlas 1558  
Db 4249 TAGAAAACTTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCCAA 4308  
QY 1558 erArgGlnGluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnS 1578  
Db 4309 GCCGTGAGGAAGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4368  
QY 1578 erLeuSerGlnSerPhePhePheLeuGluGluAlaLysHisIleAlaPheAlaIleGluLut 1598  
Db 4369 CCCTTTTCACAATCTTTTTTAACTGGAAGAGCTAAACATGCTTTTGCCATTGAAGAAAT 4428  
QY 1598 YrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluG 1618  
Db 4429 ATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTTTGTAGAACTCCTCAAGAACCAAGAG 4488  
QY 1618 luGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTyrTrpGluArgThrGlnGluA 1638  
Db 4489 AGGAGATAATAGTTGTGGAACCTTTAAACAGCACACTTTTGGTGGGAACCAACACAAGAG 4548  
QY 1638 spArgValValPhe 1642

Db 655 GATGCGCATTTATACAGTTGAAGACCAATGTTCTTTGGAAGAGCTGGATCAACT 714  
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrProArgGlyVal 220  
Db 715 AAAGCTGTTATTATGGAGAAACTGCTGTGTAGAAATAGATACCTTTCCCGGAGGATA 774  
Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
Db 775 ATTAAATATACCTAGTTATACATTTTCCCTTTTGGATATCTTTTGGCAATTCATATC 834  
Qy 241 ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
Db 835 GTAGCAGAAAAA----- 846  
Qy 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
Db 846 ----- 846  
Qy 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300  
Db 846 ----- 846  
Qy 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
Db 846 ----- 846  
Qy 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
Db 846 ----- 846  
Qy 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
Db 846 ----- 846  
Qy 361 ProPheCysHisCysThrPheValIleGlyIleAlaGln-ValMetHisLeuGluAspPh 380  
Db 847 -----GAAAGTCATGCTATTAGAGATT 870  
Qy 380 eIleGluGlyAlaSerPheSerAsnLeuThrIleGlyProTyrProLeuIleIleThrIle 400  
Db 871 TAATGAAGTGCTTCATTTTCAAAATTGACTCAGGCCCATATCCTCTAAATTATTACAA 930  
Qy 400 eIleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVa 420  
Db 931 TATCATGCTCACCTTAATAGATATTTCTATGCTCTTGGCTGTCTATCTTGATCAAGT 990  
Qy 420 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTr 440  
Db 991 CATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCCTTCATATTG 1050  
Qy 440 pSerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSe 460  
Db 1051 GTCAAAGACCAAAAGAAATATAGAGAGTTATCAGAGGCCAATGTTAATGGAAATATTAG 1110  
Qy 460 rPheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
Db 1111 TTTTAGTGAATATTATCAGCCAGTTTCTTCAGAAATTTGTAGAAAGAACCCATAGAAT 1170  
Qy 480 eSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLe 500  
Db 1171 TAGTGTATTACAGAACATACAGAAAGAGGGTGAAATGTGGAGGCTTTGAGAAATTT 1230  
Qy 500 uSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
Db 1231 GTCATTTGACATATATAGGGTTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGAAA 1290  
Qy 520 sSerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
Db 1291 GAGTACATTGATGAATATTCTTTGTGACTCTGCCACCTCTCTGATGGGTTTGCATCTAT 1350  
Qy 540 eTyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560

Db 1351 ATATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAAATGATTGGCAT 1410  
Qy 560 eCysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLe 580  
Db 1411 TTCTCCACAGTTAGATATACCTTTGATGTTTGTGACGTAGAGAAATTTTCAATTTT 1470  
Qy 580 uAlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
Db 1471 GGCTTCAATCAAAAGGATACCAGCCAAATATAATACAAAGAGTCGAGAGGTTTACT 1530  
Qy 600 uAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLeuSerGlyGlyGlnLys 620  
Db 1531 AGATTTAGACATGCAGACTATCAAGATAACCAAGCTAAATAATTAAGTGTGGTCAAAA 1590  
Qy 620 sArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspG 640  
Db 1591 AAGAAAGCTGTCTATTAGGAATTTGCTTTCTGGGAACCCAAAGATACTGCTGTAGATGA 1650  
Qy 640 uProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrAr 660  
Db 1651 ACCAAAGCTGGAAATGGACCCCTGTTCTCGATATTTGATGGAATCTTTTAAATACAG 1710  
Qy 660 GLysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAl 680  
Db 1711 AAAGCCAATCGGTCACAGTGTTCAGTACTCATTTTCATGGATGAAGTCGACATTTCTGC 1770  
Qy 680 aAspArgLysAlaValIleSerGlnClyMetLeuLysCysValGlySerSerMetPheLe 700  
Db 1771 AGATAGGAAGCTGTGATATCAAGGAATGCTGAAATGTTGGTCTTCTCAATGTTCTCT 1830  
Qy 700 uLysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaTh 720  
Db 1831 CAAAGTAATGGGGATCGCTACCGCTACGCTAGCATGTACATAGACAAATATTGTGCCAC 1890  
Qy 720 rGluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAs 740  
Db 1891 AGAATCTCTTCTCTCATCTGTTAAACAACATATACCTGGAGCTACTTTTATCAACAGAA 1950  
Qy 740 nAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPh 760  
Db 1951 TGACCAACAACCTGTGTATAGCTTGGCTTCAAGGACATGGCAAAATTTTCAGTTTGT 2010  
Qy 760 eSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrTh 780  
Db 2011 TTCTGCTCAGACAGCTCAATTCAAATTTGGTGTCATTTCTTATGTTGTTTTCATGACGAC 2070  
Qy 780 rLeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSe 800  
Db 2071 TTTGGAGAGCTATTTTAAAGCTAGAGTTGAGCAGAAATTGACCAAGCAGATTATAG 2130  
Qy 800 rValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGl 820  
Db 2131 TGTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTGATGAATGGA 2190  
Qy 820 uGlnSerLeuLeuIleLeuSerGlnThrLysAlaSerLeuValSerThrMetSerLeuTr 840  
Db 2191 ACAGAGCTTACTTATTTCTTCTGAAACCAAGGCTKCTCTAGTGAGCACCATGAGCCTTTG 2250  
Qy 840 pLysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLy 860  
Db 2251 GAAACACACAGATGTATACATACCAAGTTTCTATTTCTTACCTTGAAACCGTGAAGATA 2310  
Qy 860 sSerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPh 880  
Db 2311 ATCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTT 2370  
Qy 880 eLeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTy 900  
Db 2371 TTTGTTTCATCACCTCTTTTAAATGCTGTGGTTCCTCAACAACTGTTCCAGACTTATA 2430  
Qy 900 rPheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSe 920  
Db 2431 TTTTCTAAAACCTGGAGACAAACACATAAATAACAAACAAAGTCTGCTCTTCCAAAATTC 2490

Qy	1278	oProAspAsnGlnuAspGluAspGluAspValLysAlaGluuArgLeuLysValLysGluLe	1278
Db	3417	ACCAGACAATGAGATGAGATGAGATGCTCAAGAGCTGAAAGACTAAAGGTCAAAGAGCT	3476
Qy	1278	uMetGlyCysGlnCysCysGluGluLysProSerIleuMetValSerAsnLeuHisLysGln	1298
Db	3477	GATGGGTGCCAGTGTGTGTGAGGAGAAACCATCCATTATGGTCAGCAATTTGCATAAAGA	3536
Qy	1298	uTyTAspAspLysLysAspPheLeuLeuSerArgLysValLysLysValLalThrLysTy	1318
Db	3537	ATATGATGACAAAGAAAGATTTTCTCTTTCAAGAAAGTAAGAAAGTGGCNACTAAATA	3596
Qy	1318	rIleSerPheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGln	1338
Db	3597	CATCTCTTTCTGTGTGAAAAGGCT--ATCTTAGGACTATTGGTCCAAATGGTGTGCTGG	3654
Qy	1338	YLeSerThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPh	1358
Db	3655	CAAAAGCACAAATTAATAATTCGTGTGGTCATATTGAACCACTTCAGGCCAGGTATT	3714
Qy	1358	eLeuGlyAspTyTyrSerSerGluThrSerGluuAspAspSerLeuLysCysMetGlyTy	1378
Db	3715	TTTAGGAGATTATTTCTTCAGACACAAGTGAAGATGATGATTCACTGAAGTGTATGGGTTA	3774
Qy	1378	rCysProGlnIleAsnProLeuTyProAspThrThrLeuGlnGluHisPheGluIleTy	1398
Db	3775	CTGTCTCAGATAAACCCTTTGTGGCCAGATACATTACATTGCAGGAAACATTTTGAATTTA	3834
Qy	1398	rGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHi	1418
Db	3835	TGGAGCTGTCAAAGCAATGAGTGAAGTGAATGAAAGAACTCATAGTCAATTAACACA	3894
Qy	1418	sAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLy	1438
Db	3895	TGCATCTGATTTAAGAAGACATCTTCAGAGACTGTAAAGAAATCACTCTGCAGGAATCAA	3954
Qy	1438	sArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGln	1458
Db	3955	ACGAAAGTGTGTGTCTTAAGTATGCTAGGAATCCTCAGATTACTTTGCTAGATGA	4014
Qy	1458	uProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAla-IleArgThrA	1478
Db	4015	ACCATCTACAGGTATGGATCCCAAGCCCAACACAGCACATG-----TGCAATTCGAAC	4068
Qy	1478	laPheLysAsnArgLysArgAlaIleLeuThrThrHisTyMetGlnGluAlaGluA	1498
Db	4069	CATTATAAACAAGAGAGCGGCTGCTATPTCTGACCATCTCACTATATGGAGAGGACAGG	4128
Qy	1498	laValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValG	1518
Db	4129	CTGTCTGTGATCGAGTAGCTATCATGTGTCTGGCAGTTAAGATGTATCGACACAGTAC	4188
Qy	1518	lnHisLeuLysSerLysPheGlyLysGlyTyPheLeuGluIleLysLeuLysAspTrpI	1538
Db	4189	AACATCTAAAGAGATAAATTTGGAAGAGGCTACTTTTGGAAATTAATTTGAAGACCTGGA	4248
Qy	1538	leGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyIlePheProAsnAlas	1558
Db	4249	TAGAAACCTTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTCCCAATGCAA	4308
Qy	1558	erArgGlnGluSerPheSerSerIleLeuAlaTyLysIleProLysGluAspValGlnS	1578
Db	4309	CGCGTCAGGAAGATTTTCTTCTATTTTGGCTTATAAATTTCTTAAGAGAGATGTTCA	4368
Qy	1578	erLeuSerGlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluT	1598
Db	4369	CCCTTTCACAAATCTTTTTTAAGCTGGAGAAGCTAAACATGCTTTTGGCCATTCAGAA	4428
Qy	1598	yrSerPheSerGlnAlaThrLeuGlnGlnValPheValGluLeuThrLysGluGlnGluG	1618
Db	4429	ATAGCTTTCTCAAGCAACATGGAACAGGTTTTTGTAGAACTCAGTAAAGAAACAAGAGG	4488
Qy	1618	luGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluA	1638

Db	4489	AGGAAGATATATGTTGGAACTTTAAACAGCACACTTTGGTGGACGACGACACAGAAG	4548
Qy	1638	spArgValvalPhe	1642
Db	4549	ATAGAGTAGTATTT	4562
RESULT 14			
US-60-466-412-2441			
; Sequence 2441, Application US/60466412			
; GENERAL INFORMATION:			
; APPLICANT: CARGILL, Michele			
; APPLICANT: IAKOUBOVA, Olga			
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH			
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES TH			
; FILE REFERENCE: CL001466			
; CURRENT APPLICATION NUMBER: US/60/466,412			
; CURRENT FILING DATE: 2003-04-30			
; NUMBER OF SEQ ID NOS: 429241			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 2441			
; LENGTH: 5595			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-60-466-412-2441			
Alignment Scores:			
Pred. No.: 0 Length: 5595			
Score: 7377.00 Matches: 1474			
Percent Similarity: 89.67% Conservative: 2			
Best Local Similarity: 89.55% Mismatches: 5			
Query Match: 87.55% Indels: 167			
DB: 103 Gaps: 5			
US-10-090-458-5 (1-1642) x US-60-466-412-2441 (1-5595)			
Qy	1	MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys	20
Db	115	ATGTCCACTGCAATAGGAGGTAGGAGTTGGAGACAGACCAACACTTCTACTGAAG	174
Qy	21	AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGluIleLeuPheProLeu	40
Db	175	AATTACTTAATTAATAATCAGAACCAAAAGAGTAGTGTTCAGGAAATTCITTTTCCACTA	234
Qy	41	PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu	60
Db	235	TTTTTTTTTATTTGGTTTAATATTAAATTAGCATGATGCATCCAATAAGAATAATGAAGA	294
Qy	61	ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly	80
Db	295	GTGCCTAATATAGAACTCAATCCTATGACCAAGTTTACTCTTCTTAATCTAATTTCTTGA	354
Qy	81	TyrThrProValThrAsnIleThrSerSerIleValMetGlnLysValSerThrAspHisLeu	100
Db	355	TATACTCCAGTGACTAATATTACAAAGCAGCATCATGMAGAAAGTGTCTACTGATCATCTA	414
Qy	101	ProAspValIleIleThrGluGluTyrThrAsnGlnLysGluMetLeuThrSerSerLeu	120
Db	415	CCTGATGTCATAATTACTGAAGATATACAAATGAAAAAGAAATGTTAAATCCAGTCTC	474
Qy	121	SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg	140
Db	475	TCTAAGCCGAGCAACTTTGTAGGTGTGGTTTTTCAAGACCTCCATGTCTTATGAACATTCGT	534
Qy	141	PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer	160
Db	535	TTTTTTTCTGATGATGATCCAGTATCTTCTATTATATGGATTCAACAGCTGGCTGTTC	594
Qy	161	LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle	180
Db	595	AAATCATGTGAGGCTGCTCAGTACTGTGCTCAGTTTCAGGTTTCAGGATTCACAGCATTC	654
Qy	181	AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuThrLysGluLeuGluSerThr	200



Db 1291 GAGTACATTGATGAATATCTTTGTGGACTCTGCCACCCTTCTGTGATGGGTTGCACTCTAT 1350  
 Qy 540 eTyGlyHisArgValSerGluLeuAspGluMetPheGluAlaArgLysMetIleGlyI 560  
 Db 1351 ATATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATGGCAT 1410  
 Qy 560 cYsProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIlele 580  
 Db 1411 TTGTCCACAGTTAGATATACCTTTGATGTTTGTGACAGTAGAAGAAATTTATCAATTTT 1470  
 Qy 580 uAlaSerIleLysGlyIleProAlaAsnAsnIleleGlnGluValGlnLysValLeuLe 600  
 Db 1471 GGCTTCAATCAAGGATACCAAGCAACAATATATACAAAGAGTGCAGAGGTTTACT 1530  
 Qy 600 uAspLeuAspMetGlnThrIleLysAspGlnAlaLysLeuSerGlyGlyGlnly 620  
 Db 1531 AGATTTAGACATGCAGCATATCAAGATACCAAGCTAAAAATTAAGTGGTGGTCAAAA 1590  
 Qy 620 sArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGl 640  
 Db 1591 AAGAAAGCTGTCTATTAGGAATGCTGTTCTGGGAACCCAAAGACTGCTGCTAGATGA 1650  
 Qy 640 uProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLeuLysTyrAr 660  
 Db 1651 ACCAAGCTGGGAATGGACCCCTGTTCTCGACATATTGATGGAATCTTTTAAATACAG 1710  
 Qy 660 GLysAlaAsnAtGValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAl 680  
 Db 1711 AAAGCCATCGGTCACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATCTTGC 1770  
 Qy 680 aAspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPhele 700  
 Db 1771 AGATAGAAAGCTGTGATATCAAGAAATGCTGAAATGTGTGGTCTCTTCAATGTCTCT 1830  
 Qy 700 uLysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaTh 720  
 Db 1831 CAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCAC 1890  
 Qy 720 rGluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAs 740  
 Db 1891 AGAATCTCTTCTTCTCACTGGTAAACACATATACCTGGAGTACTTTATTACACAGAA 1950  
 Qy 740 nAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPh 760  
 Db 1951 TGACCAACAACCTTGTGTATAGCTTGGCTTTCAGGACATGGACAAATTTTCAGGTTGT 2010  
 Qy 760 eSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrTh 780  
 Db 2011 TTCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCTATTTCTTATGGTGTTCATGACGAC 2070  
 Qy 780 rLeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSe 800  
 Db 2071 TTGTGGAGAGCTATTTTAAAGCTAGAGTTGAGCAGGAAATGGACCAAGCAGATTATAG 2130  
 Qy 800 rValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGl 820  
 Db 2131 TGTATTACTACAGCAGCCTGAGAGAAATGGATTCAAAATCTTTTGTGATGAATGGA 2190  
 Qy 820 uGlnSerLeuLeuIleLeuSerGlnThrLysAlaSerLeuValSerThrMetSerLeuTr 840  
 Db 2191 ACAGAGCTTACTATTCTTCTGAAACCAAGGCTCTCTAGTGAGCAGCACCATGAGCTTGT 2250  
 Qy 840 pLysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGlySerly 860  
 Db 2251 GAAACACACAGATGTATACATAGCAAAAGTTTCATTTCTTTACCTTGAACGTTGAAGTAA 2310  
 Qy 860 sSerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPh 880  
 Db 2311 ATCAGTGAGATCAGTGTGCTCTGCTTTTAAATTTTTCAGTTTCAGATTTTATGTT 2370  
 Qy 880 eLeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTy 900  
 Db 2371 TTTGGTTTCATCACTCTTTTAAAAATGCTGTGGTTCCTCCATCAAACTGTTGCCAGCTTATA 2430

Qy 900 rPheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSe 920  
 Db 2431 TTTTCTAAAAACCTGGAGACAAACCATATAAATACAAAACAGTCTCTTCTTCAAAATTC 2490  
 Qy 920 rAla-----AspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMe 938  
 Db 2491 TGCTGGTGAGAACTCAGATATCAGTGTATTTAGTCTTTTTCACAAAGCCAGAACATAAT 2550  
 Qy 938 tValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAs 958  
 Db 2551 G----- 2551  
 Qy 958 nValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTy 978  
 Db 2552 -----GACTATGTTTTTGCAGCTGTTTTCAACAGTACTATGTTTAA 2592  
 Qy 978 rSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValTh 998  
 Db 2593 TTCTTTACCTATATTAGTGAATATCATTTAGTAATACTACTATCTTTATCAITTTAAATGTGAC 2652  
 Qy 998 rGluThrIleGlnIleTyrPheThrPhePheGlnGluIleThrAspIleValPheTy 1018  
 Db 2653 TGAACCATCCAGATCTGGAGTACCCCATTTCTTTCAAGAAATTTACTGATAGTATTTTAA 2712  
 Qy 1018 sIleGluLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTy 1038  
 Db 2713 AATTGAGCTGATTTTCAAGCAGCTTTGCTTGGAAATCATTTGTACTGCAATGCCACTTA 2772  
 Qy 1038 rPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSe 1058  
 Db 2773 CTTTGCCATCGAATAATGCAGAGATCATAGATCAAGCTTTACTCAACTTAACTTTC 2832  
 Qy 1058 rGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePh 1078  
 Db 2833 AGGTCTTTTGGCATCTGCATTTGGATTTGGAGAGCTGTTGTGATATCCCTTATTTT 2892  
 Qy 1078 eIleLeuLeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTy 1098  
 Db 2893 TATCATCTCTTATTTTGTGATAGAGCTTATTTGGCATTTTATATGGATTATATTTTA 2952  
 Qy 1098 rThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLe 1118  
 Db 2953 TACTGTAAAGTCTCTGCTGTGGTTTTTGGCTTATTTGGTTATTTGTTCCATCACTTATCT 3012  
 Qy 1118 uPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSe 1138  
 Db 3013 GTTCACCTTATATTGCTCTTTTACCTTTTAAAGAAATTTTAAATACCAAGAAATTTGGTC 3072  
 Qy 1138 rPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMe 1158  
 Db 3073 ATTTATCTATTCTGTGGCAGCGTTGGCTTGTATTGCAATCACTGAAATAAATCTTTTAT 3132  
 Qy 1158 tGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrPr 1178  
 Db 3133 GGGATACAAATTCAGAACTATCTTATATGCTTTTGTATGCTTATCATCTTCAATCATCC 3192  
 Qy 1178 oLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnVa 1198  
 Db 3193 ACTTCTAGGTGCTGATTTCTTATAGAAATTTCTTGGAGAAATGTACGAAAAAATGT 3252  
 Qy 1198 lAspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCy 1218  
 Db 3253 GGACACCTATAATCCATGGGATAGGCTTTCAGTAGCTGTTTATATCGCTTACCTCGAGTG 3312  
 Qy 1218 sValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleAr 1238  
 Db 3313 TGTACTGTGATTTTCTCTTACATACATATGAGAAAAATATATGGAGGAGCATCAATAAG 3372  
 Qy 1238 gLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluPr 1258  
 Db 3373 AAAAGATCCCTTTTTCAG-----TCTAAAAATAGGAAGCTTCCAGAAC 3416

```

Db 4429 ATAGCTTTCTCAAGCAACATGGAAACAGGTTTGTAGAACTCACTAAAGAACAGAGG 4488
Qy 1618 luGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTrpGluArgThrGlnGluA 1638
Db 4489 AGGAAGATTAAGTTGTGGAACTTTAAACAGCACACTTTGGTGGAGACCAACACAGAG 4548
Qy 1638 sPArgValValPhe 1642
Db 4549 ATAGAGTAGTATTT 4562

RESULT 13
US-60-453-135-2441
; Sequence 2441, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2441
; LENGTH: 5595
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-453-135-2441

Alignment Scores:
Pred. No.: 0 Length: 5595
Score: 7377.00 Matches: 1474
Percent Similarity: 89.67% Conservative: 2
Best Local Similarity: 89.55% Mismatches: 5
Query Match: 87.55% Indels: 167
DB: 102 Gaps: 5

US-10-090-458-5 (1-1642) x US-60-453-135-2441 (1-5595)
Qy 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20
Db 115 ATGTCCTCAATAGGAGGTAGGAGTTGGAGACAGACCAACACACTTCTACTGAAG 174
Qy 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu 40
Db 175 AATTACTTAATTAATGACAGAACCAACAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 234
Qy 41 PhePheLeuPheTrpLeuIleLysLysSerValGlnGluIleLeuPheProLeu 60
Db 235 TTTTCTTTTATTTGGTTAATTAATTAATGACATGATGATCCAAATAAGAAATATGAAGA 294
Qy 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80
Db 295 GTGCTAATATAGAACTCAATCCTATGACCAAGTTTACTCTTCTAATCTAATCTTTGGA 354
Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
Db 355 TATACTCCAGTGACTAATATTACAGCAGCATCATGAGAAAGTGTCTACTGATCATCTA 414
Qy 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
Db 415 CCTGATGTCATTAATTAATGAGATATACAAATGAAAAGAAATGTTAAATCAGTCTC 474
Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
Db 475 TCTAAGCCGAGCAACTTTGTAGGTGGTGTTCAGAGACTCCATGTCCTATGAACTTCGT 534
Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
Db 535 TTTTCTCTGATATGATTCAGTATCTCTATTTATATGATGATTCAGAGCTGGCTGTCTCA 594
Qy 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180

```

```

Db 595 AAATCATGTGAGGCTGCTCAGTACTGCTCAGGTTTTCACAGTTTTCACAGCTCCATA 654
Qy 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGlnSerThr 200
Db 655 GATGCTGCCATTTATACAGTTTGAAGCAATGTTTCTCTTGGAGAGAGCTGGAGTCACT 714
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
Db 715 AAAGCTGTTATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGATA 774
Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
Db 775 ATTTTAATATACCTAGTATACATTTTCCACTTTTGGATACTTTTGGCAATTCATATC 834
Qy 241 ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
Db 835 GTAGCAGAAAAA----- 846
Qy 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280
Db 846 ----- 846
Qy 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300
Db 846 ----- 846
Qy 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320
Db 846 ----- 846
Qy 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340
Db 846 ----- 846
Qy 341 IleGlyLeuMetIleLeuLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360
Db 846 ----- 846
Qy 361 ProPheCysHisCysThrPheValIleGlyIleAlaGln-ValMetHisLeuGluAspPh 380
Db 847 -----GAAAGTCATGCATTTAGAAAGATTT 870
Qy 380 eAsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400
Db 871 TAATGAAGGTGCTTCATTTTCAATTTGACTGCGGCCCATATCTCTAATATTATCAAT 930
Qy 400 eIleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVa 420
Db 931 TATCATGCTCACACTTAATAGTATATTCTATGTCCTGTGGCTGTCTATCTTGTATCAAGT 990
Qy 420 lIleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTr 440
Db 991 CATTCAGGAGGAATTTGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTG 1050
Qy 440 pSerLysSerLysArgAsnTyrGluGluLeuSerGluLysAsnValAsnGlyAsnIleSe 460
Db 1051 GTCAAAGAGCAAAAGAAATTTATGAGAGTTATCAGAGGCAATGTTAATGGAAATATTAG 1110
Qy 460 rPheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgI 480
Db 1111 TTTTAGTGAATTTATGAGCCAGTTTCTCAGAAATTTGTAGAAAAGAAAGCAATAGAAAT 1170
Qy 480 eSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLe 500
Db 1171 TAGTGGTATTTCAGAAAGACATACAGAAAGAGGTTGAAATTTGGAGGCTTTTGAGAAATTT 1230
Qy 500 uSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLy 520
Db 1231 GTCATTTGCATATATAGGGGTACATTTACTGCTTACTTGGCCACAGTGGACAGGAAA 1290
Qy 520 sSerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540

```

Qy 880 eLeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTy 900  
Db 2371 TTGGTTTCATCTCTTTTAAAAATGCTGTGGTTCCCATCAAACTTGTTCAGACTTATA 2430  
Qy 900 rPheLeuLysProGlyAspLysProHisLysTyLysThrSerLeuLeuLeuGlnAsnSe 920  
Db 2431 TTTTCTAAACCTCGGAGACAAACACATAAATACAAAACAAGTCTGCTTCTTCAAAATTC 2490  
Qy 920 xAla-----AspSerAspIleSerAspLeuLysSerPheThrSerGlnAsnIleMe 938  
Db 2491 TGCTGTGAGAACTCAGATATCAGTATCTTATTAGCTTTTTCACAGCCAGAAATATAT 2550  
Qy 938 tValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAs 958  
Db 2551 G----- 2551  
Qy 958 nValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTy 978  
Db 2552 -----GACTATGTTTTCAGAGCTGTTTCAACAGTACTATGTTA 2592  
Qy 978 rSerLeuProIleLeuValAsnIleLeuSerAsnTyLysLeuTyrHisLeuAsnValTh 998  
Db 2593 TTTCTTACCTATATAGTGAATATCATTAATACTACTATCTTTTATCATTTTAAATGTGAC 2652  
Qy 998 rGluThrIleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLys 1018  
Db 2653 TGAACACATCCAGATCTGAGTACCCCATCTTTCAAGAAATATCTGATATAGTTTTTAA 2712  
Qy 1018 sIleGluLeuTyPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTy 1038  
Db 2713 AATTGAGCTGTATTTTCAAGAGCTTGTGCTTGAATCATTTGTTACTGCCAATGCCACTTA 2772  
Qy 1038 rPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaValThrGlnLeuLysLeuSe 1058  
Db 2773 CTTTGGCATGGAAATGACAGAGATCATTAAGATCAAGCTTATCTCACTTAAACTTTC 2832  
Qy 1058 rGlyLeuLeuProSerAlaTyTrpIleGlyGlnAlaValAlaAspIleProLeuPhePh 1078  
Db 2833 AGGCTTTTGGCATCTGCATATTTGGATTGGACAGCTGTTGTGATATCCCTTTATTTT 2892  
Qy 1078 eIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyLysGlyLeuTyPheTy 1098  
Db 2893 TATCATTTCTTTTGTGCTAGGAAGCTTATGGCATTTTCATTATGATATATATTTTA 2952  
Qy 1098 rThrValLysPheLeuAlaValValPheCysLeuIleGlyTyValProSerValIleLe 1118  
Db 2953 TACTGTAAAGTTCTTGTCTGTGGTTTGTGCTTATTTGCTTATGTTCCATCAGTTATTTCT 3012  
Qy 1118 uPheThrTyIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSe 1138  
Db 3013 GTTCATTAATGCTTCTTTTACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTC 3072  
Qy 1138 rPheIleTySerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMe 1158  
Db 3073 ATTTATCTATTCTGTGGCAGCTTGGCTTGTATTGCAATCACTGAAATAACTTTCTTTAT 3132  
Qy 1158 tGlyTyThrIleAlaThrIleLeuHisTyAlaPheCysIleIleIleProIleTyPr 1178  
Db 3133 GGGATACAAATTTGCAACTATTCTTCAATATGCCCTTTGTATCATCATCTCCATCTATCC 3192  
Qy 1178 oLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValAtgLysAsnVa 1198  
Db 3193 ACTTCTAGGTGCTGATTTCTTTCATAAAGATTCTTGGAAAGATGTACGAAATAATGT 3252  
Qy 1198 lAspThrTyAsnProTrpAspArgLeuSerValAlaValIleSerProTyLeuGlnCy 1218  
Db 3253 GGCACCTATAATCCATGGATAGGCTTTTCAGTAGCTGTATATCGCTTACCTCAGT 3312  
Qy 1218 sValLeuTyTrpIlePheLeuLeuGlnTyTyTrpGluLysLysTyLysGlyValArgSerIleAr 1238  
Db 3313 TGTACTGTGGATTTTCTTCTTACATACTATGAGAAAAAATATGAGGAGATCAATAG 3372

Qy 1238 gLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluPr 1258  
Db 3373 AAAAGATCCCTTTTTCAG-----TCTAAAAATAGGAAGCTTCCAGAACC 3416  
Qy 1258 oProAspAsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLe 1278  
Db 3417 ACCAGACAATGAGATGAAGATGAAGATGTCAAAGCTGAAAGCTTAAAGTCAAAGCT 3476  
Qy 1278 uMetGlyCysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysG 1298  
Db 3477 GATGGGTGTCAGTGTGTGTGAGGAGAAACCATCATTTATGTCAGCAATTTGTCATAAGA 3536  
Qy 1298 uTyAspAspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTy 1318  
Db 3537 ATATGATGACAAAGAAATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATA 3596  
Qy 1318 rIleSerPheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaG 1338  
Db 3597 CATCTCTTTCTGTGTGAAAAAGGCT--ATCTTAGGACTATTGGGTCCAAATGGTGTGG 3654  
Qy 1338 yLysSerThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPh 1358  
Db 3655 CAAAGCACAAATTTAATAATTTCTGGTTGGTATATTGAACCAACTTCAGGCCAGGTATT 3714  
Qy 1358 eLeuGlyAspTy+SerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTy 1378  
Db 3715 TTTAGGAGATTATTTCTTCAGAGACAAGTGAAGATGATGATTCCTCAAGTGTATGGGTTA 3774  
Qy 1378 rCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTy 1398  
Db 3775 CTGTCTTCAGATAAACCTTTGTGGCCAGATCTACTATTCGAGGAACATTTTGAATTTA 3834  
Qy 1398 rGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHi 1418  
Db 3835 TGGAGCTGTCAAAGGAATGAGTCAAGTGACATGAAGAAGTCAATAGTCGATACACA 3894  
Qy 1418 sAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProIleGlyLys 1438  
Db 3895 TGCATCTTGAATTTAAAGACATCTTCAGAGACTGTAAAGAACTACCTGCAGGAATCAA 3954  
Qy 1438 sArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspG 1458  
Db 3955 ACNAAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCTCAGATTTACTTTGCTAGATGA 4014  
Qy 1458 uProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAla-IleArgThrA 1478  
Db 4015 ACCATCTACAGGTATGATCCCAAGCCAAACAGCAGCATG-----TGCAATTCGAAGCTG 4068  
Qy 1478 lAPheLysAsnArgLysArgAlaAlaIleLeuThrThrHisTyMetGluGluAlaGluA 1498  
Db 4069 CATTTAAAAACAGAAAGCGGCTGCTTATTTCTGCCACTCACTATATGGAGGAGGAGAGG 4128  
Qy 1498 lAValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValG 1518  
Db 4129 CTGTCTGTGATCGAGTAGCTATCATGCTGTCTGGCAGTTAAGATGTATCGGAACAGTAC 4188  
Qy 1518 lHisLysLysSerLysPheGlyLysGlyTyPheLeuGluIleLysLeuLysAspTrpI 1538  
Db 4189 AACATCTAAAGAGTAAATTTGGAAAAGGCTACTTTTGGAAATTAATTAAGAGGACTGA 4248  
Qy 1538 lGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyIlePheProAsnAlas 1558  
Db 4249 TAGAAAACCTTAGAAGTAGACCGCTTCAAAGAGAAATTCAGTATATTTTCCCAATGCAA 4308  
Qy 1558 eArgGlnGluSerPheSerSerIleLeuAlaTyLysIleProLysGluAspValGlnS 1578  
Db 4309 GCCGTGAGGAAGATTTTCTTCTTATTTTGGCTTATATAATTCCTTAGGAAGATGTTTCA 4368  
Qy 1578 eLeuSerGlnSerPhePheLeuLeuGluAlaLysHisAlaPheAlaIleGluGlu 1598  
Db 4369 CCCTTTCACATCTTTTAACTGGAAGAGCTAAACATGCTGTTTGGCATTTGAAGAT 4428  
Qy 1598 yrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluG 1618



Qy 1221 TtpIlePheLeuLeuGlnTyrTrpGluLysLysTyrGlyValGlySerIleArgLysAsp 1240  
Db 4696 TGGATTTTCTCTTCAAACTATGAGAAATAATATGGAGGCAGATCAATTAAGAAAGAT 4755  
Qy 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAsp 1260  
Db 4756 CCCTTTTTCAGAAACCTTTCAAGAGAGTCTAAATAATAGGAAGCTTCCAGAACCCAGAC 4815  
Qy 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLysValLysGluLeuMetGly 1280  
Db 4816 AATGAGATGAGATGAGATGTCAGAGCTGAAAGCTAAAGGTCAAGAGCGTATGGGT 4875  
Qy 1281 CysGlnCysGlnGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 4876 TGCACGTGTGTGAGGAGAAACCATCCATATGGTTCAGCAATTCGATTAAGAAATATGAT 4935  
Qy 1301 AspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer 1320  
Db 4936 GACAAAGAAATTTCTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 4995  
Qy 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValGlyLysSer 1340  
Db 4996 TTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGGTCCAAATGGTCTGGCAAAAGC 5055  
Qy 1341 ThrIleLeuAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 5056 ACAATTAATTAATTTCTGTTGGTGTATATGAACCACTTCAGGCCAGGTATTTTAGGA 5115  
Qy 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 5116 GATTATTTCTTCAGAACCAAGTGAAGATGATGATTCCTCAAGTGTATGGGTACTGTCT 5175  
Qy 1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
Db 5176 CAGATAAACCTTTGTGGCCAGATCTATCATTCGAGCAACATTTTGAATTTATGGAGCT 5235  
Qy 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db 5236 GTCAAGGAATGAGTGCAGTGCATGAAAGAGTCAATAGTCAATTAACACATGCACCT 5295  
Qy 1421 AspLeuLysGluHisLeuGlnLysThrValLysLeuProAlaGlyIleLysArgLys 1440  
Db 5296 GATTTAAGAAACATCTTCAGAGAGCTGTAAGAAACTACCTGCAGAAATCAAGCAAG 5355  
Qy 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
Db 5356 TTGTGTTTCTCTTAAGTATGCTAGGGAATCCTCAGATTAATTTGTAGATGAACCATCT 5415  
Qy 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
Db 5416 ACAGGTATGATCCCAAGCCAAACAGCACATGTGCGAGCAATTCGAATCGATTTAAA 5475  
Qy 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluAlaGluAlaValCys 1500  
Db 5476 AACGAAAGCGGGCTGCTATTTCTGACCCTCATATATGGAGAGGCGAGGGCTGCTCT 5535  
Qy 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
Db 5536 GATCGAGTAGTATCATGGTGTCTGGGCGAGTAAAGATGATCGGAACAGTACAACTCA 5595  
Qy 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLysLysAspTrpIleGluAsn 1540  
Db 5596 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGGCTGGATAGAAAC 5655  
Qy 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
Db 5656 CTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCCGCTCAG 5715  
Qy 1561 GluSerPheSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
Db 5716 GAAAGTTTTTCTTATTTTGGCTTATAAATTTCCCTAAGGAAGATGTTCACTCCCTTCA 5775  
Qy 1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600

Db 5776 CAATCTTTTTTAAGCTGGAAGAACTAAACATGCTTTTGCCATTGAAGATATAGCTTT 5835  
Qy 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluAsp 1620  
Db 5836 TCTCAAGCAACATTTGGAACAGAGTCTTCTAGAACTCACTAAAGAACAGAGGAGGAGAT 5895  
Qy 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640  
Db 5896 AATAGTTGTGGAACCTTTAAACACACACTTTGGTGGGAACGAAACACAGAAAGATAGTA 5955  
Qy 1641 ValPhe 1642  
Db 5956 GTATTT 5961  
RESULT 12  
US-60-453-050-2441  
; Sequence 2441, Application US/60453050  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele  
; APPLICANT: LUKE, May  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CUG01457  
; CURRENT APPLICATION NUMBER: US/60/453,050  
; CURRENT FILING DATE: 2003-03-10  
; NUMBER OF SEQ ID NOS: 82762  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2441  
; LENGTH: 5595  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-60-453-050-2441  
Alignment Scores:  
Pcted. No.: 0 Length: 5595  
Score: 7377.00 Matches: 1474  
Percent Similarity: 89.67% Conservative: 2  
Best Local Similarity: 89.55% Mismatches: 5  
Query Match: 87.55% Indels: 167  
DB: 102 Gaps: 5  
US-10-090-458-5 (1-1642) x US-60-453-050-2441 (1-5595)  
Qy 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
Db 115 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCAGAACACTTCTACTGAAG 174  
Qy 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGluIleLeuPheProLeu 40  
Db 175 AATTACTTAATTAATGCAAGAACCAAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 234  
Qy 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
Db 235 TTTTATTTATTTGGTTAATTAATTAATTAGCATGATGATGATCCAAATAAGAAATATGAAGAA 294  
Qy 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80  
Db 295 GTGCTTAATATAGAACCTCAATCCCTATGCAAGATTTACTCTTTCTAATCTAATCTTGGGA 354  
Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
Db 355 TATACTCAGTGCATTAATATTAACAAGCAGCATCATGAGAAAGTGTCTACTGATCATCTA 414  
Qy 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
Db 415 CCTGATGTCATTAATCTGAAGAAATACAAATGAAAAGAAATGTTTAACATCCAGCTC 474  
Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
Db 475 TCTAAGCCGAGCAACTTTGTAGGTGTGTTTTCAGAGACTCCATGCTCTATGAATTCGT 534  
Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160

2678 AGTGGTATTTCAGAGACATACAGAAAGAGAGGTTGAAATGTGGAGGCTTTGAGAAATTTG 2737  
 501 SerPheAspIleTyrGlnGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 2738 TCATTGACATATAGAGGGTCAGATTACTGCTTACTTGGCCACAGTGGACAGAAAG 2797  
 521 SerThrLeuMetAsnIleLeuGlyGlyLeuGlyCysProSerAspGlyPheAlaSerIle 540  
 2798 AGTACATTGATGAATATCTTTGTGGACTCTCCACCTCTCGATGGGTTTGCATCTATA 2857  
 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 2858 TATGGACACAGAGCTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATGGCATT 2917  
 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGlnLeuSerIleLeu 580  
 2918 TGTCCACAGTTAGATATACATTGATGTTTTCAGCAGTAGAAGAAATTTATCAATTTTG 2977  
 581 AlaSerIleGlyGlyLeuProAlaAsnAsnIleLeuGlnValGlnLysValLeuLeu 600  
 2978 GCTTCATCAAGAGGATACACCCCAACATATAATACAGAGAGTGCAGAGGTTTACTA 3037  
 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 3038 GATTAGACATGCAGACTATCAAGATACCAAGCTTAAATAATTAAGTGGTGGTCAAAA 3097  
 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
 3098 AGAAGCTGCTCATTAGGAATGCTGTTCTTGGAAACCCCAAGATGCTGCTGATAGAA 3157  
 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyrArg 660  
 3158 CCAACAGCTGGAATGACCCCTGTTCTGCACATATTGATGGAATCTTTTAAATPACAGA 3217  
 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 3218 AAGCCCAATCGGGTGACGTGTCAGTACTCATTTTCATGGATGAAGCTGACATTTTCA 3277  
 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 3278 GATAGGAAGCTGTGATATCACAGGAATGCTGAAATGTTGGTCTTCAATGTTCTCTC 3337  
 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 3338 AAAAGTAAATGGGGATCGGCTACCGCTGACATGATGATAGACAAATATTGTGCCACA 3397  
 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyValAlaThrLeuGlnGlnAsn 740  
 3398 GAATCTCTTCTCAGTGGTTAAACACACATATACCTGGAGCTACTTTATTACACAGAT 3457  
 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
 3458 GACCAACAATTTGTATAGCTTGCCTTTCAAGGACATGACAAATTTTCAGGTTTGTGT 3517  
 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
 3518 TCTGCCCTTAGACATGATTTCAAAATTTGGGTGTCATTTCTTATGGTGTTCATGACGACT 3577  
 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
 3578 TTGGAAGACGTATTTTAAAGCTAGAAATTTGAAGCAGAAATTTGACCAAGCAGG----- 3630  
 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
 3630 ----- 3630  
 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
 3630 ----- 3630  
 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
 3630 ----- 3630

861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
 3631 -----TTGCTTCTGCTTTTAAATTTTTCACAGTTTTCAGATTTTATGTTT 3675  
 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
 3676 TTGGTTCATCATCTCTTTTAAATGCTGTGGTCCCATCAAACTTGTTCAGAGCTTATAT 3735  
 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
 3736 TTTCTAAACCTCGAGACAAACACATATAATACAAACCAAGTCTGCTTCTTCAAAATCT 3795  
 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
 3796 GCTGACTCAGATATCAGTATCTTATAGCTTTTTCAGAGCCAGCAATATATGGTACG 3855  
 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
 3856 ATGATTAATGACAGTGACTATGTATCGTGGCTCCCATAGTGGCGCTTAAATGTGGTG 3915  
 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
 3916 CATTCAAGAAAGGACTATGTTTTCAGAGCTGTTTTCACAGTACTATGGTATTCTTTA 3975  
 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
 3976 CTTATATTAGTGAATATCACTAGTACTATCTTTTATCATTTAAATGTGACTGAAACC 4035  
 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
 4036 ATCCAGATCTGGAGTACCCCATTTCTCAAGAAATTTACTGATATAGTTTAAATGTAG 4095  
 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAla 1040  
 4096 CTGTATTTTCAAGCAGCTTTGCTTGGATCATTTGTTACTGCAATGCCACCTTACTTTGCC 4155  
 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
 4156 ATGGAATATCAGAGATCATATAAGATCAAGCTTTACTCAACTTAAACTTTTCAAGTCTT 4215  
 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
 4216 TTGCCATCTGCATTTGGATGGATGGCAAGCTGTTGTTGATATCCCTTATTTTATCAT 4275  
 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
 4276 CTTATTTTGTATGCTAGGAGCTTATTTGGCATTTTCACTATGATTTATTTTATCTGTA 4335  
 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
 4336 AAGTTCCTTGTGTGGTTTTGTGCTTATGGTATGTTCCATCAGTTATTTCTGTTCACT 4395  
 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
 4396 TATATGCTCTTCTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATATC 4455  
 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
 4456 TATTCTGTGCAGCGTTGGCTTGTATTGCAATCACTGAAATAACTTTCTTTATGGGATAC 4515  
 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
 4516 ACAATGCAACTATTTCTTCACTTATGCTTTTGTATCATCATCTCAATCTCACTTCTTA 4575  
 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
 4576 GGTTCCTGATTTCTTTTCAATAAGATTTCTTGGAGAATGTACGAAAAATTTGGGACACC 4635  
 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
 4636 TATATCCATGGGATAGGCTTTTCACTAGCTTTTATTCGCTTACCTGACGTGTGTACTG 4695

```

APPLICANT: JACKSON, Alan A.
APPLICANT: JIANG, Xin
APPLICANT: JIN, Pei
APPLICANT: KABLE, Amy E.
APPLICANT: KHARE, Reena
APPLICANT: LEE, Soo Y.
APPLICANT: LEE, Sally
APPLICANT: MARQUIS, Joseph P.
APPLICANT: RAMKUMAR, Javalaxmi
APPLICANT: RICHARDSON, Thomas W.
APPLICANT: SWARNARAR, Anita
APPLICANT: TRAN, Uyen K.
APPLICANT: CHAWLA, Narinder K.
APPLICANT: WILSON, Amy D.
TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
FILE REFERENCE: PF-1435 PCT
CURRENT APPLICATION NUMBER: PCT/US03/14026
CURRENT FILING DATE: 2003-05-02
PRIOR APPLICATION NUMBER: US 60/377,444
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: US 60/377,435
PRIOR FILING DATE: 2002-05-03
PRIOR APPLICATION NUMBER: US 60/386,497
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: US 60/388,180
PRIOR FILING DATE: 2002-06-11
NUMBER OF SEQ ID NOS: 132
SOFTWARE: PERL Program
SEQ ID NO 75
LENGTH: 6170
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7511310CB1
PCT-US03-14026-75

Alignment Scores:
Pred. No.: 0 Length: 6170
Score: 8028.50 Matches: 1573
Percent Similarity: 95.86% Conservative: 1
Best Local Similarity: 95.80% Mismatches: 68
Query Match: 95.28% Indels: 1
DB: 1 Gaps: 1

US-10-090-458-5 (1-1642) x PCT-US03-14026-75 (1-6170)
QY 1 MetSerThrAlaIleArgGluValTrpArgGlnThrArgThrLeuLeuLeuLys 20
Db 1238 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACAGACACTTCTACTGAAG 1297
QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGlnLeuLeuPheProLeu 40
Db 1298 AATTACTTAATTAATGCGACCAAAAGAGTAGTTTCAGGAATTTCTTTCCACTA 1357
QY 41 PhePheLeuPheTrpLeuIleLeuLeuSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 1358 TTTTTTTTATTTGGTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGAT 1417
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80
Db 1418 GTCCCTTAATAGAACTCAATCTATGACAGAGTTTACTTCTTCTTAATCTAATTTCTTGA 1477
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisIleu 100
Db 1478 TATACTCCAGTACTAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 1537
QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
Db 1538 CCTGATGTCATAATTTACTGAAGAAATATACAAATGAAGAAATATGTTTACATCCAGTCTC 1597
QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140

```

```

1598 TCTAAGCCGAGCAACTTTGTAGGTGTGTTTCAAGAGCTCCATGCTCTATGAACCTTGGT 1657
141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
1658 TTTTTCCTGATATGATCCAGTATCTTCTATTTATATGATTTCAAGAGCTGGCTGTTCAT 1717
161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180
1718 AAATCATGTGAGGCTGCTCAGTACTGGTCCCTCAGGTTTTCACAGTTTTCACAGCATCCATA 1777
181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200
1778 GATGCTGCCATTTATACAGTTGAGACCAATGTTCTCTTTGGAGAGAGCTGGAGTCAACT 1837
201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
1838 AAAGCTGTTTATTATGGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGATTA 1897
221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
1898 AITTTTAATATACCTAGTTATAGCATTTTACCTTTTGGATACCTTTTGGCAATTTCATATC 1957
241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
1958 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATATATGGGACTTCATGATACT 2017
261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280
2018 GCCTTTTGGCTTTCCCTGGGTTCTTCTATATACAGTTTAAATTTTCTATGTCCTCTCT 2077
281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300
2078 ATGGCAGTCATTCGACAGCTTCTTTGTTTATTTCTCTAAAGTAGCAGCATTTGTGATATT 2137
301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320
2138 CTGCTTTTCTTCTTATGATTTATCATCTGATTTTCTTCTTATGCTGACACCTCTT 2197
321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340
2198 TTTTAAAAATCAAAACATGTTGGGAAATAGTTGAATTTTGTACTGTGGCTTTGGATTT 2257
341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360
2258 ATTGGCTTTATGATATCTCTCATAGAAAGTTTCCCAATCGTTAGTGTGGCTTTTCAGT 2317
361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380
2318 CCTTTCTGTCACGTACTTTTGTGATTTGATTTGACAGTCAATGATGATGATGATGATGAT 2377
381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLysLeuIleThrIle 400
2378 AATGAAGGTCTTCTCATTTTCAAAATTTGACTGCGAGGCCATATCTCTTAATTTACAAAT 2437
401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420
2438 ATCATGCTCACTTAATATAGTATTTCTATGCTCTCTTGGCTGCTCTATCTTGTATCAAGTC 2497
421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440
2498 ATTCCAGGGGAAATTTGGCTTACGAGATCATCTTATTTATTTTCTGAAGCCCTTCATATTG 2557
441 SerLysSerLysArgAsnTyrGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460
2558 TCAAAGACAAAAGAAATATGAGGAGTTATCAGAGGCCAATGTTTAAATGGAATATATAGT 2617
461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480
2618 TTTAGTGAATTTATGAGCCAGTTTCTTTCAGATTTGTAGAAAGAGAGCAATGAAT 2677
481 SerGlyIleGlnLysThrTyrArgLysLysGlyValGluAsnValGluAlaLeuArgAsnLeu 500

```



941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaLeuAerValMet 960  
2821 ATGATTAATGACAGTACTATGATTCCTGCTCCCATAGTGGCGCTTTAAATGTGRTG 2880  
961 HisSerGluLysAspTyrValPheAlaValPheAsnSerThrMetValTyrSerLeu 980  
2881 CATTGAGAAAGGACTATGTTTTCGACGTGTTTCAACAGTACTATGTTTTCCTTTA 2940  
981 ProIleLeuValAsnIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
2941 CCTATATTAGTGAATATCATTTAGTAACACTATCTTTATCATCTTTAAATGTGACTGAAACC 3000  
1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
3001 ATCCAGATCTGGAGTACCCCTCTTTTCAAGAAATTTACTGATATAGTTTAAATTTGAG 3060  
1021 LeuTyrPheGlnAlaLeuLeuGlyIleValThrAlaMetProProTyrPheAla 1040  
3061 CTGTAATTTCAAGCAGCTTCTTGGATCATTTGTTACTGCAATGCCACCTTACTTTGCC 3120  
1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
3121 ATGGAAATGTCAGAGATCATTAAGATCAAGCTTAYACTCAACTTAACTTTTCAGGTCTT 3180  
1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleIle 1080  
3181 TTGCCATCTGCATATTGGATTGGACACGCTGTTGTTGATATCCCTTATTTTATCATTT 3240  
1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
3241 CTATATTTGATGCTAGGAAGCTTATTTGCAATTTCAATTATGGATTATATTTTATATCTGA 3300  
1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
3301 AAGTTCCTTGTGCTGCTTTTTCCTCTTATGTTGTTATGTTTCCATCAGTTATTTCTGTTCACT 3360  
1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
3361 TATATGCTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
3421 TATTTCTGGGACGCTGCTGTTGATTCGATCACTGAAATAACTTTCTTTATGGGATAC 3480  
1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
3481 ACAATTGCAACTATTTCTTCATTTATGCTTTTGTATCATCTCAATCTATCCACTTCTA 3540  
1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
3541 GGTTCCTGATTTCTTTTCAATAAGATTTCTTGGAAAGATGTACGAAATAATGTGGACACC 3600  
1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
3601 TATATCCATGGGATAGGCTTTTCACTAGCTGTTATATGCTTACCTGACGTGTGACTG 3660  
1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
3661 TGGATTTTCTCTTCACTATCATGAGAAATAATGAGGAGCAGATCAATTAAGAAAGAT 3720  
1241 ProPhePheAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
3721 CCTTTTTCAGAAACCTTTCAACGAGATCTTAAATATAGAACTTCCAGAACCCAGAC 3780  
1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
3781 AATGAGGATGAAGATGAAGATGTCAAGCTGAAAGACTAAAGGTCAAGAGAGCTGATGGGT 3840  
1281 CysGlnCysCysGluGluLysProIleMetValSerAsnLeuHisGlyLeuTyrAsp 1300  
3841 TGCCAGTGTGTGAGAGAGAAACCATCTTATGTTGTCAGCAATTTGCATAAGAAATATGAT 3900  
1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320

3901 GACAAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCACTTAATATCATCTCT 3960  
1321 PheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValaGlyLysSer 1340  
3961 TTTCTGTGTGAAAGAGGAGAGATCTTAGGACTATTGGGTCCAAATGGTGTGGCAAAAGC 4020  
1341 ThrIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
4021 ACAATTAATTAATTTCTGTTGTTGATATTGAACCAACTTCAGGCCAGATATTTTAGGA 4080  
1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
4081 GATTATTTCTCAGACCAAGTGAAGATGATTCATCTGAAGTGTATGGTTACTGTCTCT 4140  
1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
4141 CAGATAAACCCCTTTTGTGGCCAGATATCATTTGCGAGCAACTTTTGAATTTTATGGAGCT 4200  
1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
4201 GTCAAAGGAATGAGTCCAAAGTCACTGAAGAAGTCAATGATGCAATAACACATGCACTT 4260  
1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
4261 GATTTTAAAGAACATCTTCAGAAGACTGTAAAGAAACTACCTGCAGGAATCAACCGAAAG 4320  
1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
4321 TTGTGTTTGTCTCTAAGTATGCTAGGGAATCTTCAGATTACTTTGCTAGATGACCATCT 4380  
1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
4381 ACAGGTATGGATCCCAAGCCAAACAGACACATGTGGCAGCAATTCGAATCGAATTTAA 4440  
1481 AsnArgLysArgAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCys 1500  
4441 AACAGAAAGCGGCTGCTATCTGACCCTACTATATGAGGAGGAGCAGGCTGTCTGT 4500  
1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
4501 GATCGAGTAGCTATCATGTGTGCTGGCAGTTAAGATGTATCGGAACAGTCAACATCTA 4560  
1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
4561 AAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTTAAATTTGAAGACTGGGATAGAAAC 4620  
1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
4621 CTAGAAGTAGACCCCTTCAAGAGAAATTCAGTATATTTTCCCAATCCAGCCGTGAG 4680  
1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
4681 GAAAGTTTTTCTTCTTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTCA 4740  
1581 GlnSerPhePheLysLeuGluGlu 1588  
4741 CAATCTTTTTTAAAGCTGGAAGAA 4764

RESULT 11  
PCT-US03-14026-75  
; Sequence 75 Application PC/TUS0314026  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE CORPORATION  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: BECHA, Shanya D.  
; APPLICANT: BULLOCH, Sean A.  
; APPLICANT: CHANG, Hei-Ru  
; APPLICANT: ELLIOTT, Vicki S.  
; APPLICANT: EMERLING, Brooke M.  
; APPLICANT: GRIFFIN, Jennifer A.  
; APPLICANT: HAFALIA, April J.A.  
; APPLICANT: ISON, Craig H.

601 AAAGCTGTTATTATGGAGAAAGCTGCTGTTGTAGAAATAGATACCTTTCCCGAGAGTA 660  
 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 661 ATTTTAAATACCTAGCTATAGCATTTTCCACCTTTTGGATACCTTTTGGCAATTCATATC 720  
 241 ValAlaGluGlySerGlyIleValGlyPheLeuIleValMetGlyLeuHisAspThr 260  
 721 GTAGCAGAAAAAGAAAAATATAAGAAATTTTAAAGATATGGGATTCATGATACT 780  
 261 AlaPheTyrLeuSerTyrValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 781 GCCTTTTGGCTTCTCGGGTCTCTCTATAWACAAGTTTAAATTTCTTATGTCCTTCTT 840  
 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300  
 841 ATGGCAGTCATTCGACAGCTCTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTATATT 900  
 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 901 CTGCTTTTCTTATGGATATCATCTGTATTTTGTGCTTTTAAATGCTGACACCTCTT 960  
 321 PheLeuGlySerIleHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 961 TTTAAAAATCAAAACATGTGGATAGTTGAATTTTGTGCTTTTGTGCTTTTGGATTT 1020  
 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTyrLeuPheSer 360  
 1021 ATTTGGCTTATGATATCTCATAGAAAGTTTCCCAATCTGTAGTGTGGCTTTTCACT 1080  
 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 1081 CCTTTCTGTCACGTGATCTTTGTGATTTGTGATTTGACAGGTTCATGCAATTTAGAAATTT 1140  
 381 AsnGluGlyAlaSerPheSerLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 1141 AATGAAGTGCTTCATTTTCAAAATTTGACTGAGGCCCATATCTCTAAATTAATTAATTT 1200  
 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 1201 ATCATGTCACACTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTATCAAGTC 1260  
 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuIleProSerTyrTrp 440  
 1261 ATTCAGGGGAAATTTGGCTTTACGGATCATCTTTATATTTTCTGAAGCTTCATATTGG 1320  
 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 1321 TCAAGAGCAAAAGAAATATGAGGATTTATCAGAGGGCAATGTTAAATGGAATATTAGT 1380  
 461 PheSerGluIleIleGluProValSerSerGluPheValGlyIleValAlaIleArgIle 480  
 1381 TTTAGTGAATATTATGAGCCAGTTCTTCAGAAATTTGTAGGAAAAAGAGCCATAGAATT 1440  
 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
 1441 AGTGGTATTTCAGAGACATACAGAAAGAGGGTGAATGTCGAGGCTTTGAGAAATTTG 1500  
 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 1501 TCATTTGACATATATGAGGGTCAGATTACTGCTTACTTGGCCACAGTGGCAACAGAAAG 1560  
 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
 1561 AGTACATTTGATGATATTTCTTTGTGGACTCTGCCCACCTTCTGATGGGTTTGCATCTATA 1620  
 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 1621 TATGGACACAGAGCTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATTTGGCAT 1680  
 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 1681 TGTCACAGTTAGATATACCTTTGATGTTTTGACAGTAGAAGAAAAATTTATCAATTTTG 1740

QY 581 AlaserIleGlyGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 Db 1741 GCUTCAATCAAGGATACCCAGCCCAATATATAAAGAGTGAGAGGTTTTACTA 1800  
 QY 601 AspleuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 Db 1801 GATTTAGCATGACAGCTATCAAGATNACCAAGCTTAAANAATTAAGTGGTGGTCAAAA 1860  
 QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
 Db 1861 AGAAAGCTGTCATTAGGAATGCTGTTCTTGGGAACCAAGATACCTGCTGCTAGATGAA 1920  
 QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLeuLysTyrArg 660  
 Db 1921 CCACACCTGGAATGGACCCCTGTTCTCGACATATTGTAATGTAATCTTTTAAATACAGA 1980  
 QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 Db 1981 AAAGCCAATCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATCTCTGCA 2040  
 QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 Db 2041 GATAGGAAGCTGTGATATCACAAAGGAATGCTGAATGTGTTGTTCTTCAATGTTCCCTC 2100  
 QY 701 LysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 Db 2101 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATACACAAATATTGTGCCACA 2160  
 QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
 Db 2161 GAATCTCTTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATTACAAAGAAAT 2220  
 QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
 Db 2221 GACCAACAATCTGTGTATAGCTTGCCTTTCAAGGACATGGACAAATTTTCAGSTTTGTT 2280  
 QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
 Db 2281 TCTGCCCTAGACAGCTATTCAAAATTTGGGTGGCATTTCTTATGGGTTTCCATGAGCACT 2340  
 QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
 Db 2341 TTGGAAACGCTATTTTAAAGCTAGAGTTGAGGAGAAATTTGACCAAGCAGATTATAGT 2400  
 QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
 Db 2401 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTATGAAATGGAA 2460  
 QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
 Db 2461 CAGAGCTTACTTATCTTTCTGAAACCAAGGCTKCTCTAGTGAGCACCATGAGCCTTTGG 2520  
 QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
 Db 2521 AAACAACAGATGATACAAATAGCAAGTTTCAATTTCTTACCTTGAACCTGAAAGTAAA 2580  
 QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
 Db 2581 TCAGTGTAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTTCAGATTTTATGTTT 2640  
 QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
 Db 2641 TTGGTTTCACTCTTTTAAAAATGCTGTGTTTCCCATCAAACTGTTCACAGCTTATAT 2700  
 QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
 Db 2701 TTTCTAAAAACCTGGAGCAAAACACATAAATACAAACCAAGTCTGCTTCTTCAAAATCT 2760  
 QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
 Db 2761 GCTGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAAGCCAGAACATATAATGGTGACG 2820

QY	1295	uHlslyeGluTyrrAspAspLysLysAaspPheLeuLeuSerArgLysValLysValAl	1315
Db	3919	GCATAAGAATATGATCAACAAGAAGATTTCTCTCTTCAGAAACCAAGAAGTGGC	3978
QY	1315	aThrLysTyrrIleSerPheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAs	1335
Db	3979	AACTAATAATACATCTCTTCTCTGTGMAAAAGGAGAGATCTTAGCACTATTGGGTCCAA	4038
QY	1335	nGlyValAGlyLysSerThrIleIleAenIleLeuValGlyAspIleGluProThrSerGI	1355
Db	4039	TGGTGGCTGGCAAAAGCAAAATTATTAATATTCTGGTGGTGATATTGAACCACTTCAGG	4098
QY	1355	yGlnValPheLeuGlyAspTyrrSerSerGluThrSerGluAspAspSerLeuLysCy	1375
Db	4099	CCAGGTATTTTTAGGAGATTATCTTCAGAGACAAGTGAAGATGATCATCTCACTGAAGTG	4158
QY	1375	sMetGlyTyrcysProGlnIleAenProLeuTyrrProAspThrThrLeuGlnGluHisph	1395
Db	4159	TATCGGTACTGTCCTCAGATAAACCCCTTGTGTGCCAGATACTACATTGAGGAACATTT	4218
QY	1395	eGluIleTyrrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerAr	1415
Db	4219	TGAAATTTATGGAGCTGTCAAAGGAATGAGTGCAAGTGACATGAAGAAGATCATAGTCG	4278
QY	1415	gIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAl	1435
Db	4279	AATAACACATGCACCTGATTATAAAGAACATCTTCAGAAGACTGTAAAGAAGTACCTCGC	4338
QY	1435	aGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrIle	1455
Db	4339	AGGAATCAACAGAAAGTTGTGTTTGTCTTAAGTATGCTAGGGNATCCTCAGATTACTTT	4398
QY	1455	uLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIle	1475
Db	4399	GCTAGATGAACACATCTCAGATATGGATGCCAAAGCCAAAAGCACATGTGGCGAGCAAT	4458
QY	1475	eArgThrAlaPheLysAenArgLysArgAlaAlaIleLeuThrThrHisTyrrMetGluGI	1495
Db	4459	TCGAACTGCAATTTAAACACAAAGCGGGCTGCTATTCTGCACCACTCACATATGGAGGA	4518
QY	1495	uAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGI	1515
Db	4519	GGCAGAGGTTGTCTGTGATCGAGTAGCTATCATGTGTCTGGGCAGTTAAGATGTATCGG	4578
QY	1515	yThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrrPheLeuGluIleLysLeuLy	1535
Db	4579	AACAGTACACATCTAAAGAGTAAATTTGAAAGGGCTACTTTTGGAAATTAATTTGAA	4638
QY	1535	sAspTrpIleGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrrIlePhePr	1555
Db	4639	GGACTGGATAGAAAACCTTAGAAGTAGACCCGCTTCAAAGAGAAATTCAGTATATTTCC	4698
QY	1555	aAsnAlaSerArgGlnGlnSerPheSerSerIleLeuAlaTyrrLysIleProLysGluAs	1575
Db	4699	AAATGCAAGCCGTGAGAAAAGTTTTCCTCTATTCTTATTTGGCTTATATAAATTCCTAAGGA	4758
QY	1575	pValGlnSerLeuSerGlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIle	1595
Db	4759	TGTTCACTCCCTTTACAACTCTTTTTTAAAGCTGGAAAGAGCTAAACATGCTTTTGCCAT	4818
QY	1595	eGluGluTyrrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGI	1615
Db	4819	TGAAGATATAGCTTTTCTCAAGCAACATTTGGAACACAGGTTTTTTGTAGAACTCACTAAAGA	4878
QY	1615	uGlnGluGluGluAspAsnSerCysGlyThrLeuAenSerThrLeuTyrrTrpGluArgThr	1635
Db	4879	ACAAGAGGAGGAATATAGTTGTGGAACTTTAAACAGCACACTTTGGTGGAGACGAAC	4938
QY	1635	rGlnGluAspArgValPhe	1642
Db	4939	ACAAGAGATAGAGTAGTAGTATTTT	4960

RESULT 10

US-09-971-121-3

US-09-971-121-3

Sequence 3, Application US/09971121

GENERAL INFORMATION:

APPLICANT: Hu, Yi

APPLICANT: Nepomnich, Boris

TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding t

TITLE OF INVENTION: Same

FILE REFERENCE: LEX-0250-USA

CURRENT APPLICATION NUMBER: US/09/971,121

CURRENT FILING DATE: 2001-10-04

PRIOR APPLICATION NUMBER: US 60/239,629

PRIOR FILING DATE: 2000-10-10

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 4785

TYPE: DNA

ORGANISM: homo sapiens

US-09-971-121-3

Alignment Scores:

Pred. No.: 0 Length: 4785

Score: 8112.00 Matches: 1583

Percent Similarity: 99.69% Conservative: 0

Best Local Similarity: 99.69% Mismatches: 5

Query Match: 96.27% Indels: 0

DB: 41 Gaps: 0

US-10-090-458-5 (1-1642) x US-09-971-121-3 (1-4785)

Qy 1 MetSerThrAlaileArgGluValGlyValTrArgGlnThrArgThrLeuLeuLeuLys 20

Db 1 ATGTCCACTGCAATTAGGGAGGTAGGAGTTGGAGACAGACCAGACACTTCTACTGAAG 60

Qy 21 AsnTyrLeuileLysCysAsnThrLysLysSerSerValGlnGluileLeuPheProleu 40

Db 61 AATTACTTAATAAATGTCAGAACCAAAAGAGTAGTGTTCCAGGAAATCTCTTTTCCACTA 120

Qy 41 PhePheLeuPheTrpLeuileLeuileSerMetMetHisProAsnLysLysTyrGluGlu 60

Db 121 TTTTITTTATTTGGTTAATATTAATAGCATGATGCATCCAAATAGAATAATATGAGAA 180

Qy 61 ValProAsnileGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuileLeuGly 80

Db 181 GTGCCTAATATAGAACTCAATCTCATGGACAAGTTTACTCTTCTAATCTAATCTTGGA 240

Qy 81 TyrThrProValThrAsnileThrSerSerIleMetClnLysValSerThrAspHisLeu 100

Db 241 TATACCTCCAGTACTAATATTAACAGCAGCATCATGCAGAAAGTGTTCTACTGATCATCTA 300

Qy 101 ProAspValileileThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120

Db 301 CCTGATGTCATAATTACTGAAGAAATACAAATGMAAAGAAATGTTAATCATCCAGTCTC 360

Qy 121 SerLysProSerAsnPhelValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140

Db 361 TCTAAGCCGAGCAACTTTGTAGTGTTGGTTTTCAGAGACTCCCATGTCTATGAATTCGT 420

Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160

Db 421 TTTTITTCCTGATATGATTCAGTATCTTCTATTTATATGGATTCAGAGCTGCTGTTC 480

Qy 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerile 180

Db 481 AAATCATGTGAGCGCTCTCAGTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540

Qy 181 AspAlaAlaileileGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200

Db 541 GATGCTGCCATATACAGTTGAGACCAATGTTTCTTTTGGAGGAGCTGGAGTCAACT 600

Qy 201 LysAlaValileMetGlyGluThrAlaValValGluileAspThrPheProArgGlyVal 220



; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS

; FILE REFERENCE: PF-1591 PCT

; CURRENT APPLICATION NUMBER: PCT/US03/33087

; CURRENT FILING DATE: 2003-10-16

; PRIOR APPLICATION NUMBER: US 60/419,313

; PRIOR FILING DATE: 2002-10-16

; PRIOR APPLICATION NUMBER: US 60/421,033

; PRIOR FILING DATE: 2002-10-23

; PRIOR APPLICATION NUMBER: US 60/421,349

; PRIOR FILING DATE: 2002-10-25

; PRIOR APPLICATION NUMBER: US 60/423,516

; PRIOR FILING DATE: 2002-10-04

; NUMBER OF SEQ ID NOS: 112

; SOFTWARE: PERL Program

; SEQ ID NO 82

; LENGTH: 5004

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc\_feature

; OTHER INFORMATION: Incyte ID No: 7523469CBI

PCT-US03-33087-82

Alignment Scores:

Pred. No.:	0	Length:	5004
Score:	8253.50	Matches:	1618
Percent Similarity:	98.50%	Conservative:	7
Best Local Similarity:	98.18%	Mismatches:	7
Query Match:	97.95%	Indels:	17
DB:	1	Gaps:	2

US-10-090-458-5 (1-1642) x PCT-US03-33087-82 (1-5004)

Qy	1	MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys	20
Db	50	ATGCCAAGTCAATAGGAGGTAGGAGTTGGAGACAGACAGACCACTTCTACTGAAG	109
Qy	21	AsnTyrLeuLeuLysCysArgThrLysLysSerValGlnGluLeuLeuPheProLeu	40
Db	110	AAATCTTAATTAATGACAGAACCAAAAGAGTAGTGTTCAGGAATTCCTTCCACTA	169
Qy	41	PhePheLeuPheTrpLeuLeuLeuLeuSerMetMetHisProAsnLysLysTyrGluGlu	60
Db	170	TTTTTTTTTATTTGGTTAAATTAATTAATGATGATGATGATGATGATGATGATGATG	229
Qy	61	ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLeu	80
Db	230	GTGCTTAATATAGAACCAATCTTATGACAAAGTTTACTCTTCTAATCTAATCTTGG	289
Qy	81	TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu	100
Db	290	TATCTCCAGTGACTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	349
Qy	101	ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu	120
Db	350	CCTGATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	409
Qy	121	SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg	140
Db	410	TCTAAGCCGACCACTTTGAGGTGTGTTTCAAGACCTCCATGCTCTATGAACTTCGT	469
Qy	141	PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer	160
Db	470	TTTTTCTGATATGATTCAGTATCTTCTATTTATATGATGATGATGATGATGATGATG	529
Qy	161	LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle	180
Db	530	AAATCATGTGAGGTGCTCAGTACTGCTCAGGTTTCAGGTTTCAGGTTTCAGGTTTC	589
Qy	181	AspAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr	200
Db	590	GATGCTGCCATTATACAGTTGAAGCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACT	649

Qy	201	LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal	220
Db	650	AAAGCTGTATTATGGGAGAACTCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA	709
Qy	221	IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle	240
Db	710	ATTTAATATACCTAGTTTATAGCATTTTCACCTTTGGATACCTTTTGGCAATTCATATC	769
Qy	241	ValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr	260
Db	770	GTAGCAGAAAAGAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATCTACT	829
Qy	261	AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu	280
Db	830	GGCTTTTGGCTTTCCTGCTTCTTATATACAGTTTAAATTTTCTTATGCTCCCTCTT	889
Qy	281	MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe	300
Db	890	ATGGCAGTCATTCGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTGATATT	949
Qy	301	LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu	320
Db	950	CTGCTTTTCTCTTATGAGTATCATCTGTATTTTCTGCTTAAATGCTGACACTCTT	1009
Qy	321	PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe	340
Db	1010	TTTAAAAAATCAAAACATGTTGGAATTTTGTACTGTGGCTTTTGGATTT	1069
Qy	341	IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer	360
Db	1070	ATTGGCTTATGATTAATCTCATAGAAAGTTTCCCAATCGTTAGTGTGGCTTCCAGT	1129
Qy	361	ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe	380
Db	1130	CCCTTTCTGCTACTGACTTTTGTGATTTGTTATGTCACAGGTCATGCATTTTGAAGATTT	1189
Qy	381	AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle	400
Db	1190	AATGAGGTGCTTCTTCTTCAATTTGACTCGAGCCCATATCTCTAATTTATTAACAATT	1249
Qy	401	IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal	420
Db	1250	ATCATGCTCACACTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTGATCAAGTC	1309
Qy	421	IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp	440
Db	1310	ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCCTCATATTGG	1369
Qy	441	SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer	460
Db	1370	TCAAAGAGCAAAAGAAATTTATGAGGAGTTATCAGAGGGCAATGTTAATGGAATATCAGT	1429
Qy	461	PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle	480
Db	1430	TTTAGTGAATTTATGAGCCAGTTTCTTCAGAAATTTGAGGAGGCTTTGAGAAATTTG	1489
Qy	481	SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu	500
Db	1490	ACTGTTATTCAGAACACATACAGAAAGGGGTGAAATGTGGAGGCTTTGAGAAATTTG	1549
Qy	501	SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys	520
Db	1550	TCATTTGACATATATGAGGTCAGATTACTGCTTACTTGGCCACAGTGGCAACGAGAAAG	1609
Qy	521	SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle	540
Db	1610	AGTACATTTGATGAATATTTCTTTTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA	1669
Qy	541	TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle	560
Db	1670	TATGGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAATGATTGGCATT	1729
Qy	561	CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu	580



Db 1014 CTGCTTTTTCCTTTATGATATATCATCTGTATTTTGTCTTTAATGCTGACACCTCTT 1073  
Qy PheLysLysSerLysHisValGlyLleValGluPhePheValThrValAlaPheGlyPhe 340  
Db 1074 TTTAAAAAATCAAAACATGCGGAAATAGTGGAAATTTTGTACTGTGGCTTTGGATTT 1133  
Qy IleGlyLeuMetIleIleLeuIleGlySerPheProLysSerLeuValTrpLeuPheSer 360  
Db 1134 ATTGGCTTTATGATAATCCTCATGAAAGTTTCCCAATCGTTAGTGTGGCTTTTCAGT 1193  
Qy ProPheCysHisCysThrPheValIleGlyLleAlaGlnValMetHisLeuGluAspPhe 380  
Db 1194 CCTTCTCTACTGTACTTTGTGATTTGGTATTTGCACAGGTCATGCAATTTAGAAATTTT 1253  
Qy AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
Db 1254 AATGAGGTGCTTCATTTTCAAAATTTGACTGCGGCCCATATCTCTAAATTTATTAAT 1313  
Qy IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTrpLeuAspGlnVal 420  
Db 1314 ATCATGCTCACACTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTTGATCAAGTC 1373  
Qy IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
Db 1374 ATTCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCATATTGG 1433  
Qy SerLysSerLysArgAsnTyrGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
Db 1434 TCAAGAGCAAAAGAAATATGAGAGTTATCAGAGGCAATGTATATGGAATATTTAGT 1493  
Qy PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
Db 1494 TTTAGTGAATATTTGACCCAGTTTCTTCAGATTTTGTAGAAAGAACCCATTAAGATTT 1553  
Qy SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
Db 1554 AGTGTATTCAGAACACATACAGAAAGAGGTGAAATGTGGAGCTTTTGAGAAATTTG 1613  
Qy SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
Db 1614 TCATTTGACATATATGAGGGTCAGATTTACTGCTTACTTGGCCACAGTGGAAACGAAAG 1673  
Qy SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
Db 1674 AGTACATGATGAATATTTTGTGACTTGCACCTTCTGATGGTTTTCATCTATA 1733  
Qy TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db 1734 TATGACACAGAGTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAATGATTTGGCAT 1793  
Qy CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
Db 1794 TGTCACAGTTAGATATACACTTTGATGTTTGTGACAGTAGAAGAAATTTATCAATTTG 1853  
Qy AlaSerIleLysGlyIleProAlaAsnIleIleGlnGluValGlnLysValLeuLeu 600  
Db 1854 GCTTCAATCAAGGATACCAACCAATATATATACAAAGATGCGAAGGTTTACTA 1913  
Qy AsnLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
Db 1914 GATTTAGATGTCAGACTATCAAGATTAACAGAGTAAATAATTAAGTGGTGGTCAAAA 1973  
Qy ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
Db 1974 AGAAAGCTGTCAATAGGAATTTGCTGTTCTTGGGAACCCCAAGATACTGCTGTAGTAA 2033  
Qy ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyrArg 660  
Db 2034 CCAACAGCTGGAAATGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAATACAGA 2093  
Qy LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680

Db 2094 AAAGCCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTTGC 2153  
Qy AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2154 GATAGAAAGCTGTGATATCACAGGAATGCTGAAATGTGTGGTTCTTCAATGTCTCTC 2213  
Qy LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2214 AAAAGTAAATGGGATCGCTACCGCTGAGCATGTATACATAGCAAAATATTGTGCCACA 2273  
Qy GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db 2274 GAATCTCTTCTTCTACTGGTTAAACAACATATACCTGGAGCTACTTTATTACACAGAAT 2333  
Qy AsnGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 2334 GACCAACAATCTGTGTATAGCTTGGCTTTCAAGGACATGGCAAAATTTTCAGGTTGTTT 2393  
Qy SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2394 TCTGCCCTAGACAGTCATTTCAAAATTTGGGTGTCAATTTCTTATGGTGTTCATGACGACT 2453  
Qy LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2454 TTGGAAGAGCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTCACCAAGCAGATTTAGT 2513  
Qy ValPheThrGlnGlnProLeuGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2514 GTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAATCTTTTGATGAATGGAA 2573  
Qy GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 2574 CAGAGCTTACTTATTTCTGAAACCAAGGCTGCTAGTGAGCACCATGAGCCCTTTGG 2633  
Qy LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db 2634 AAACAACAGATGATACAAATAGCAGAGTTTCATTTTCTTTACCTTGAACGTAAGATAA 2693  
Qy SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2694 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTT 2753  
Qy LeuValHisHisSerPheLysAsnAlaValValProLysLeuValProAspLeuTyr 900  
Db 2754 TTGGTTTCATCCTCTTTTAAATGCTGGTTCCTCATCAAACTTGTTCACAGCTTATAT 2813  
Qy PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
Db 2814 TTTCTAAAACCTGGAGACAAACCAATAAATACAAACAAAGTCTGCTTCTTCAAAATTC 2873  
Qy Ala-----AspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
Db 2874 GCTGGTGAGAGTGTNNGTGAAGACTCAGATATCAGTGATCTTTATTAGCTTTTTCACAGC 2933  
Qy GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
Db 2934 CAGAAATATGTTGACGATGATTAATGACAGTACTATGATCCGTGCTCCCAATAGT 2993  
Qy AlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSer 974  
Db 2994 CGGGCTTTAAATGTGATGCTTTCAGAAAAGGACTATGTTTTCGAGCTGTTTTCACAGT 3053  
Qy ThrMetValTyrSerLeuProLysValAsnIleLeuSerAsnTyrTyrLeuTyrHis 994  
Db 3054 ACTATGGTTTATTTTACCTATATTAGTGAATATCATTTAGTAACCTACTTATTATCAT 3113  
Qy LeuAsnValThrGluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAsp 1014  
Db 3114 TTAATGATGACTGAAACCATCCAGATCTGGAGTACCCCATCTTTCAGAAATTTACTGAT 3173  
Qy IleValPheLysIleGluLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAla 1034  
Db 3174 ATAGTTTTTAAATGAGCTGTATTTTCAAGCAGCTTTTCTTGGAAATCATTTGTAACCTGC 3233



1394 sPheGluIleTyrGlyAlaValIysGlyMetSerAlaSerAspMetLysGluValIleSe 1414  
4170 TTTTGAAATTTATGGAGCTGTCAAAGGAATGAGTCAAGTGCATGAAGAAGTCAATAG 4229  
1414 rArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuPr 1434  
4230 TCGAATAACACATGCACCTTGATTTAAAGAACATCTTCAGAAGACTGTAAAGAACTACC 4289  
1434 oAlaGlyIleLysArgLysLeuLysCysPheAlaLeuSerMetLysGluValIleTh 1454  
4290 TCGAGGAATCAACACGAAAGTGTGTTGCTCTAAGTATGCTAGGAAATCCTCAGATTAC 4349  
1454 rLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTTPArgAl 1474  
4350 TTTGCTAGATGAACCATCTACAGGTATGATCCCAAGCCAAACAGCACATGTGGCGAGC 4409  
1474 aIleArgThrAlaPheLysAsnArgLysArgAlaLalLeuThrThrHisTyrMetGl 1494  
4410 AATTTCGAACCTGATTTAAAGAACAGAAAGCGGCTGCTATTCGACCACTCACTATATGGA 4469  
1494 uGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIl 1514  
4470 GAGGCGAGAGGCTGCTGTGATCGAGTAGCTATCATGGTGTCTGGCGAGTAAAGATGTAT 4529  
1514 eGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLe 1534  
4530 CGAACAGATACAAACATCTAAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAAT 4589  
1534 uLysAspTTPLeGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePh 1554  
4590 GAAGGACTGGATAGAAACCTAGAGTAGACCGCCTTCAAGAGAAATTCAGTATATTTT 4649  
1554 eProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAlaTyrLysIleProLysGl 1574  
4650 CCCAAATGCAAGCGCTCAGAGAGTATTTTCTCTATTTTGGCTTATAAAATTCCTAAGGA 4709  
1574 uAspValGlnSerLeuSerClnSerPheLysLysLeuGluAlaLysHisAlaPheAl 1594  
4710 AGATGTTTCAGTCCCTTTCACAACTTTTATAGCTGGGAAGAGCTAAACATGCTTTTGC 4769  
1594 aIleGluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLy 1614  
4770 CATTTGAAGAATATAGCTTTTCTCAAGCAACATTTGGAACAGGTTTGTAGAACTCACTAA 4829  
1614 sGluGlnGluGluGluAsnSerCysGlyThrLeuAsnSerThrLeuTyrTTPGluAr 1634  
4830 AGAACAGAGAGGAGAGATAATAGTTGTGGAACCTTTAAACAGCACACTTTGTGGGAACG 4889  
1634 gThrGlnGluAspArgValPhe 1642  
4890 AACACAAGAGATAGATAGTATTT 4914

RESULT 8  
US-10-090-458-1  
; Sequence 1, Application US/10090458  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Hongyun  
; APPLICANT: Kilinski, Ligia  
; APPLICANT: Le Bihan, Stephanie  
; TITLE OF INVENTION: NOVEL RECS TRANSPORTER AND USES THEREOF  
; FILE REFERENCE: 100103.403  
; CURRENT APPLICATION NUMBER: US/10/090,458  
; CURRENT FILING DATE: 2002-03-01  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 5463  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 25, 2888, 2889  
; OTHER INFORMATION: n = A, T, C or G

US-10-090-458-1  
Alignment Scores:  
Pred. No.: 0 Length: 5463  
Score: 8309.00 Matches: 1629  
Percent Similarity: 98.85% Conservative: 1  
Best Local Similarity: 98.79% Mismatches: 2  
Query Match: 98.61% Indels: 18  
DB: 44 Gaps: 2  
US-10-090-458-5 (1-1642) x US-10-090-458-1 (1-5463)  
Qy 1 MetSerThrAlaIleArgGluValGlyValTTPArgGlnThrArgThrLeuLeuLys 20  
Db 114 ATGTCCACTGCAATATTAGGAGGTAGGAGTTTGGAGACAGACAGACACTTCTACTGAAG 173  
Qy 21 AsnTyrIleLysCysArgThrLysLysSerValGlnGluLeuPheProLeu 40  
Db 174 AATTACTTAATTAATTCAGAACCAAAAGAGTAGTGTTCAGAAATTTCTTTTCCACTA 233  
Qy 41 PhePheLeuPheTTPLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
Db 234 TTTTITTTTATTTTGGTTAATATTAAATTAGCATGATGCATCCAAATAAGAAATATGAAGA 293  
Qy 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLys 80  
Db 294 GTGCCTAATATAGAACTCAATCCTATGGACAAAGTTTACTCTTTCTTAATCTAATCTTGTGA 353  
Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
Db 354 TATACTCCAGTGACTAATATTACAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA 413  
Qy 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
Db 414 CCTGATGTCATAATTAATCTGAAGAATATACAAATGAAAGAAATGTTTAACTCCAGCTCTC 473  
Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
Db 474 TCTAAGCCGAGCAACTTTGTAGTGTGTTTCAAGAGACTCCCATGCTCTTCTGACTTCGT 533  
Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
Db 534 TTTTTCCTCATATGATTCCTCAGTATCTTCTATTTATATGATTTCAAGAGCTGGCTGTCA 593  
Qy 161 LysSerCysGluAlaIleGlnTyrTTPSerSerGlyPheThrValLeuGlnAlaSerIle 180  
Db 594 AAATCATGTGAGGCTGCTCAGTACTGCTCCTCAGGTTTTCAGTTTTCAGCATCCATA 653  
Qy 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTTPLysGluLeuGluSerThr 200  
Db 654 GATGCTGCCATTATACAGTTGAAGACCAATGTTTCTCTTTGGAAGAGAGCTGGAGTCAACT 713  
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
Db 714 AAAGCTGTTATATGGAGAACTGCTGTTGTAGAATAAGATACCTTTTCCCGAGAGATA 773  
Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
Db 774 ATTTTAAATACCTAGTTATAGCATTTTTCACCTTTTGGTACTTTTGGCAATTCATATC 833  
Qy 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260  
Db 834 GTAGCAGAAAAAGAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATGATACT 893  
Qy 261 AlaPheTTPLeuSerTTPValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
Db 894 GCCTTTTGGCTTTCCTGGGTCTCTCTATATACAGTTTAAATTTTCTTATGTCCTTCTT 953  
Qy 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
Db 954 ATGGCAGTCATTCGACAGCTTCTTTGTTATTTTCTCAAGAGTAGCAGCATTTGTGATATT 1013  
Qy 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320

DB 1981 AAAGCCAAATCGGTCAGAGTCTCAGTACTCAATTTTCATGGATGAAGCTGACATTTCTTGC 2040  
QY AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
DB 2041 GATAGGAAGCTGTGATATCACAGGAATGCTGAATGTGTGTGTCTTCAATGTTCTC 2100  
QY LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 2101 AAAAGTAAATGGGGATCGGTCACCGCTGAGCATGTACATAGACAAATATTTGCCACA 2160  
QY GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
DB 2161 GAACTCTCTTCTCACTGGTTAAACAACATATACCTGGAGCTACTTTATTACACAGAT 2220  
QY AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
DB 2221 GACCAACAACATGTGTATAGTGTGCTTTTCAAGGACATGACAAATTTTCAGGTTTCTTT 2280  
QY SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 2281 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCTTTCTTATGGTGTCTTCCATGACACT 2340  
QY LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
DB 2341 TTGGAAGACGATATTTTAAAGCTAGAGCTTGAAGCAGAAATTTGACCAAGCAGATTAAGT 2400  
QY ValPheThrGlnGlnProLeuGluGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 2401 GTATTTCTCAGCAGCCACTGGAGAAAGATGATTCAAATCTTTTGATGAATGGAA 2460  
QY GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
DB 2461 CAGAGCTTACTATTTCTTCTGAAACCAAGCTCTCTAGTGAGCACCATGAGCTTTGG 2520  
QY LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
DB 2521 AAAACAACAGATGTATACAAATAGCAAGATTCATTTCTTTACCTTGAACACGTGAAGTAA 2580  
QY SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 2591 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTT 2640  
QY LeuValHisHisSerPheLysAsnAlaValProIleLysLeuValProAspLeuTyr 900  
DB 2641 TTGGTTCATCTCTTTTAAATATGCTGTGGTTCCTCCATCAAACTTTGTCAGACTATAT 2700  
QY PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
DB 2701 TTTCTAAACCTGGAGCAACCAACATTAATACAAACAGCTGCTCTTCTTCAAAATCT 2760  
QY Ala-----AspSerAspIleSerAspLeuIleSerPheThrSer 934  
DB 2761 CTTGGTGAGAGTGTNNGTGAAGACTCAGATATCATGATCTTATTAGCTTTTTCACAAAGC 2820  
QY GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
DB 2821 CAGAACATATGGTGAGATGATTAATGACAGTACTATGTATCCGTGGCTCCCATAGT 2880  
QY AlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSer 974  
DB 2881 CCGGCTTTAAATGTGATGCTATTCAGAAAGGACTATGTTTTCAGCTGTTTTCACAGT 2940  
QY ThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHis 994  
DB 2941 ACTATGGTTTATCTTTTACCTTATTTAGTGAATATCATATTAGTAACTACTATCTTTATCAT 3000  
QY LeuAsnValThrGluThrIleGlnIleTrpSerThrProPheGlnGluIleThrAsp 1014  
DB 3001 TTAATGTGACTGAACCACTCCAGATCTGGAGTACCCCATCTTTTCAAGAAATTAATCAT 3060  
QY IleValPheLysIleGluLeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAl 1034  
DB 3061 ATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATATTGAACTGC 3120

QY 1034 aMetProProTyrPheAlaMetGluAsnAlaGluAenHisLysIleLysAlaTyrThrGln 1054  
DB 3121 AATGCCACCTACTCTGCCATCGAAAATGACAGAAATCATAA----- 3162  
QY 1054 nLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValAspIle 1074  
DB 3163 -----GGTCTTTTTCCTCTCTGCATATGATGGATGGACAGCTGTTGTTGATAT 3209  
QY 1074 eProLeuPhePheIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrGln 1094  
DB 3210 CCCCATTATTTTATCATCTTATTTTGTATGCTAGGAAGCTTATTTGGCATTTCAATATG 3269  
QY 1094 yLeuTyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValPr 1114  
DB 3270 ATTATATTTTATCTGTAAGTTCCTTCTGCTGTGTTTTTTCCTTATTTGGTATTTGCC 3329  
QY 1114 oSerValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLys 1134  
DB 3330 ATCAGTTATTTCTGTCATCTATATTTCTTTCACCTTTAAGAAATTTTAAATACCA 3389  
QY 1134 sGluPheTrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIle 1154  
DB 3390 AGAATTTTGGTTCATTTATCTATCTGTGCGAGCTTGGCTTGTATTTGCAATCACTGAAAT 3449  
QY 1154 eThrPhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleTrl 1174  
DB 3450 AACTTTCTTTTATGGGATACCAATTTGCAACTATTTCTTCAATATGCTTTTGTATCATCAT 3509  
QY 1174 eProIleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnVa 1194  
DB 3510 TCCAACTATTCACCTTCTAGGTTGCTGATTTCTTTTCATAAAGATTTCTTGGAGAAATGT 3569  
QY 1194 lArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSerPr 1214  
DB 3570 ACGAAAAATGCGACACCTATAATCCATGGATAGGCTTTTCAGTAGCTGTATATCCGCC 3629  
QY 1214 oTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyGln 1234  
DB 3630 TTACCTGCGAGTGTGACTGTGGATTTTCTCTTCAATACATATGAGAAAAATATGGAGG 3689  
QY 1234 yArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLys 1254  
DB 3690 CAGATCAATAAGAAAGATCCCTTTTTCAGAAACCTTTTCAACGAAGTCTTAAATAATAGGAA 3749  
QY 1254 sLeuProGluProProAspAsnGluAspGluAspValLysAlaGluAtrLeuLys 1274  
DB 3750 GTTTCAGAACCCACGACATGAGATGAGATGAGATGTCAAGCTGAAAGCTAAA 3809  
QY 1274 sValLysGluLeuMetGlyCysGlnCysGlnGluLysProSerIleMetValSerAs 1294  
DB 3810 GGTCAAGAGCTGATGGGTTGCGAGTGTGTGAGGAGAAACCATCCATTTATGTCAGCAA 3869  
QY 1294 nLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSerArgLysValLysLysVa 1314  
DB 3870 TTTGCAATAAGAAATATGATGACAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAT 3929  
QY 1314 lAlaThrLysTyrIleSerPheCysValLysGlyGluIleLeuGlyLeuLeuGlyPr 1334  
DB 3930 GGCACATAAATACATCTCTTCTGTGTCAAAAAAGAGAGATCTTTAGGACATTTGGGTCC 3989  
QY 1334 oAsnGlyAlaGlyLysSerThrIleIleAsnIleLeuValGlyAspIleGluProThrSe 1354  
DB 3990 AAATGGTGTGCAAAAGCAAAATTTATTAATTTCTGTTGTTGATATTTGAACCAACTTC 4049  
QY 1354 rGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAspSerLeuLys 1374  
DB 4050 AGGCCAGGATTTTATGAGATTTATTTCTTCAGAGCAAGTGAAGATGATGATTTCACTGAA 4109  
QY 1374 sCysMetGlyTyrCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHi 1394  
DB 4110 GTGTATGGGTTACTGCTCTCCTAGATAAACCTTTTGGCCAGATACTATCATTTGCGAGGAA 4169

US-10-090-458-3

Alignment Scores:

Pred. No.:	0	Length:	4917
Score:	8309.00	Matches:	1629
Percent Similarity:	98.85%	Conservative:	1
Best Local Similarity:	98.79%	Mismatches:	2
Query Match:	99.61%	Indels:	18
DB:	44	Gaps:	2

US-10-090-458-5 (1-1642) x US-10-090-458-3 (1-4917)

QY	1	MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys	20
DB	1	ATGTCCACGTCAATTTAGGAGGTAGGAGTTTGGACAGACACAGAACACTTCTACTGAAG	60
QY	21	AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu	40
DB	61	AATTACTTAATTAATGCAGACCAAAAGAAGTAGTAGTTTCAAGAAATCTTTTTCCACTA	120
QY	41	PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu	60
DB	121	TTTTTTTTTATTTGGTTAATAATTAAGCATGATGCATCCAAATAAGAATAATGAAGAA	180
QY	61	ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly	80
DB	181	GTGCCTAATATAGAACTCAATCCTATGACCAAGTTTACTCTTTCTAATCTAATCTTTGA	240
QY	81	TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu	100
DB	241	TATACTCCAGTGACTAATATTACAAGCAGCATCATGCAGAAAGTGTCTACTGATCATCTA	300
QY	101	ProAspValIleIleThrCluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu	120
DB	301	CCGTGATGTCATTAATTACTGAAGATATACAATGAAGAAGAATGTTAAACATCCAGTCTC	360
QY	121	SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg	140
DB	361	TCTAAGCCGAGCAACTTTGTAGTGTGGTTTCAAAGACTCCATGCTCTATGAATTCGT	420
QY	141	PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer	160
DB	421	TTTTTTTCCGTGATATGATCCAGTACTCTTCTATTTATATGGAATCAAGAGCTGGCTGTCA	480
QY	161	LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle	180
DB	481	AAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTCACAGTTTACAGCATCCATA	540
QY	181	AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr	200
DB	541	GATGTGCCATTTATACAGTTGAAGACCAATGTTTCTCTTTGGAGGAGCTGGAGTCAACT	600
QY	201	LysAlaValIleMetGlyCluThrAlaValValGluIleAspThrPheProArgGlyVal	220
DB	601	AAAGCTGTTATTATGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA	660
QY	221	IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle	240
DB	661	ATTTTAATACCTAGTTATAGCATTTTCACCTTTTGGTACATTTTGGCAA'TTCATATC	720
QY	241	ValAlaGluLysGluLysLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr	260
DB	721	GTAGCAGAAAAGAGAAAAAATAAAGAAATTTTAAAGATATATGGGACTTTCAATATCT	780
QY	261	AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu	280
DB	781	GCCTTTTGGCTTTCTCGGGTCTTCTATATACAAAGTTTAATTTTTCTATGTCCCTCTT	840
QY	281	MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe	300
DB	841	ATGGCAGTCATTCGACAGCTCTTTGTATTCTCTCAAGTAGACAGCATTTGATATT	900
QY	301	LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu	320

901	CTGCTTTTTTCCCTTTATAGGATTTATCATCTGTATTTTTTGTCTTTAATGCTGACACCTCTT	950
321	PhcLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe	340
961	TTTAAAAAAATCAAAACATGTGGGAATAGTTGAATTTTTTGTCTACTGTGGCTTTTGGATTT	1020
341	IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTprLeuPheSer	360
1021	ATTGGCCCTTATGATATATCTCATAGAAAGTTTTTCCCAATCGTTAGTGTGGCTTTTCAGT	1080
361	ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMethHisLeuGluAspPhe	380
1081	CCCTTCTGTCACATGTACTTTTTGTGATGTGTTATGTCACAGGTCAATGATTTAGAAGATTTT	1140
381	AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle	400
1141	AATGAAGGTGCTTCATTTTCAAAATTTGACTGTCAGGCCCATATCTCTAATTAATCAATT	1200
401	IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal	420
1201	ATCATGCTCACACTTAATAGTATATTCTATGTCTCTTGGCTGTCTATCTTGATCAAGTC	1260
421	IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp	440
1261	ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTCTGAAGCCCTTCATATTGG	1320
441	SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer	460
1321	TCAAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGGCAATGTTAATGGAAATATTAGT	1380
461	PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle	480
1381	TTTAGTGAATTAATTGAGCCAGTTCTTCAGAAATTTGTAGGAAAGAGCCATAGAAATT	1440
481	SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu	500
1441	AGTGGTATTGAGAAGACATACAGAAAGAGGGTGAAATGTGGAGGCTTTGAGAAATTTG	1500
501	SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys	520
1501	TCATTTGACATATATGAGGGTCAGATTACTGCTTACTTTGGCCACAGTGGACAGGAAG	1560
521	SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle	540
1561	AGTACATTGATGAATATCTTTGTGGACTTCGCCACCTTCGTATGGGTTTGCATCTATA	1620
541	TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle	560
1621	TATGGACACAGACTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATAATGATGGCATT	1680
561	CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu	580
1681	TGTCCACAGTTAGATATACACTTTGATGTTTTCACAGTAGAAGAAATTTATCAATTTTG	1740
581	AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu	600
1741	GCITCAATCAAAGGGATACCAGCAACAATATAATAACAAGAAAGTGCAGAAGGTTTTTACTA	1800
601	AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys	620
1801	GATTTAGACATGCAGACTATCAAAAGATAACCAAGCTTAAAAAATTTAAGTGGTGTCAAAA	1860
621	ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu	640
1861	AGAAAGCTGTCATTAGGAATTGCTGTTCTTTGGGAACCCCAAGATACTGCTGTAGATGA	1920
641	ProThrAlaGlyMetAspProCysSerArgHisIleValTprAsnLeuLeuLysTyrArg	660
1921	CCAACAGCTGGAAATGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAAAATACAGA	1980
661	LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAlaSpIleLeuAla	680

1041 MetGluAsnAlaGluAsnHisIysIleLysAlaThrGlnLeuLysLeuSerGlyLeu 1060  
1061 LeuProSerAlaThrTrpIleGlyGlnAlaValAlaSpIleProLeuPhePheIleIle 1080  
1101 LysPheLeuAlaValPheCysLeuIleGlyThrValProSerValIleLeuPheThr 1120  
1121 TyrIleAlaSerPheThrPheLysIleLysAlaPheHisThrGlyLeuThrPheThrVal 1140  
1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyThr 1160  
1161 ThrIleAlaThrIleLeuHisThrAlaPheCysIleIleIleProIleThrProLeuLeu 1180  
1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProThrLeuGlnCysValLeu 1220  
1221 TrpIlePheLeuLeuGlnThrGlyLysIleThrGlyArgSerIleArgLysAsp 1240  
1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
1261 AsnGluAspGluAspValLysAlaGluArgLeuLysValLysValLysGluLeuMetGly 1280  
1281 CysGlnCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluThrAsp 1300  
1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysThrIleSer 1320  
1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValaGlyLysSer 1340  
1341 ThrIleLeuAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
1361 AspThrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyThrCysPro 1380  
1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleThrGlyVala 1400  
1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420

3375 ATGAAAGAAATGAGTCAAGTCAATGAAAGAAAGTCAATGAAAGTCAATGAAAGTCAATG 4514  
1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
1451 GATTTAAAGAAACATCTTCAGAAAGACTGTAAAGAACTACTGCGAGGAATCAAAACGAAG 4574  
1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
1461 ThrClyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
1481 AsnArgLysArgAlaAlaIleLeuThrThrHisThrMetGluGluAlaGluAlaValCys 1500  
1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
1521 LysSerLysPheGlyLysGlyThrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
1541 LeuGluValAspArgLeuGlnArgGluIleGlnThrIlePheProAsnAlaSerArgGln 1560  
1561 GluSerPheSerSerIleLeuAlaThrLysIleProLysGluAspValGlnSerLeuSer 1580  
1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluLysThrSerPhe 1600  
1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpGluArgThrGlnGluAspArgVal 1640  
1641 ValPhe 1642  
5175 GTATT 5180

## RESULT 7

US-10-090-458-3  
Sequence 3, Application US/10090458  
GENERAL INFORMATION:  
APPLICANT: Chen, Hongyun  
APPLICANT: Killinski, Ligia  
APPLICANT: Le Bihan, Stephane  
TITLE OF INVENTION: NOVEL ABCAS TRANSPORTER AND USES THEREOF  
FILE REFERENCE: 100103.403  
CURRENT APPLICATION NUMBER: US/10/090,458  
NUMBER OF SEQ ID NOS: 5  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 4917  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 2775, 2776  
OTHER INFORMATION: n = A,T,C or G

Db 1155 CTGCTTTTTCCTTTATGGATTATCATCTGTATTTTTCCTTTAATGCTGACACCTCTT 1214  
Qy PheLysLysSerLysHisValGlyIleValGluPheValThrValAlaPheGlyPhe 340  
Db TTTAAAAAATCAAAACATGTGGGAATAGTGAATTTTGTGTACTGTGGCTTTTGGGATTT 1274  
Qy IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
Db ATTGGCTTATGATAATCCTCATAGAAAGTTTCCCAAATCGTTAGTGGCTTTTCAGT 1334  
Qy ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
Db CTTTCTCTCACTGTACTTTTGTGATGGTATTTGACAGGTCATGGCAATTTAGAGAATTTT 1394  
Qy AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProFyrProLeuIleIleThrIle 400  
Db RATGAAGGTGCTTCATTTTCAAAATTGACTGCAGGCCCATATCTCTAAATTTACAAATT 1454  
Qy IleMetLeuThrLeuAsnSerIlePheFyrValLeuLeuAlaValTrpLeuAspGlnVal 420  
Db ATCATGCTCACACTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTTGATCAAGTC 1514  
Qy IleProGlyGluPheGlyLeuArgSerSerLeuTrpPheLeuLysProSerTrp 440  
Db ATTCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCTTCATATGG 1574  
Qy SerLysSerLysArgAsnFyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
Db TCAAGAGACAAAGAAATTTATGAGAGATTATCAGAGGCAATGTTAATGGAAATATTAGT 1634  
Qy PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
Db TTTAGTGAATTTATTGACCAAGTTTCTTCAGAAATTTGTAGAAAGAACCCNTAAGAATT 1694  
Qy SerGlyIleGlnLysThrFyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
Db AGTGGTATTTCAGAGACATACAGAAAGAGGTGAATAATGTGGAGCTTTTGAGAAATTTG 1754  
Qy SerPheAspIleFyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
Db TCATTTGACATATATGAGGTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGGAAG 1814  
Qy SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540  
Db AGTACATTGATGAATATTCTTTGTGGACTCTGCCACCTCTCGATGGGTTTGCATCTATA 1874  
Qy TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db TATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAAAATGATTGGCATT 1934  
Qy CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
Db TGTCCACAGTTAGATATACACTTTGATGTTTTCAGTAGAAGAAATTTATCAATTTTG 1994  
Qy AlaSerIleLysGlyIleProAlaAsnIleIleGlnGluValGlnLysValLeuLeu 600  
Db GCTTCAATCAAGGATACCAAGCCCAATATAATACAGAAAGTGCAGAGGTTTACTA 2054  
Qy AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
Db GATTTAGACATGCAGACTATCAAGATTAACCAAGCTAAAAATTAAGTGGTGGTCAAAA 2114  
Qy ArgLysLeuSerLeuGlyIleAlaValLeuGluGlyAsnProLysIleLeuLeuLeuAspGlu 640  
Db AGAAAGCTGTCTATTAGAAATTTGCTTCTTGGGAACCCAAAGATACTGCTGTAGATGAA 2174  
Qy ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTrpArg 660  
Db CCAACAGCTGGATGGACCCCTGTCTCGACATATTGTATGGAAATCTTTTAAAAATACAGA 2234  
Qy LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680

Db 2235 AAAGCCAATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTGCA 2294  
Qy AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db GATAGAAAGCTGTGATATCACAGGAATGCTGAATGTGTGGTTCTTCAATGTCTCTC 2354  
Qy LysSerLysTrpGlyIleGlyIleGlyIleArgLeuSerMetTrpIleAspLysTrpCysAlaThr 720  
Db AAAAGTAAATGGGGATCGCTACCGCTGAGCATGTATACATAGACAAAATATTGTGCCACA 2414  
Qy GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
Db GAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2474  
Qy AspGlnGlnLeuValTrpSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db GACCAACAATCTGTGTATAGTCTTCTTCAAGGACATGGACAAATTTTCAGGTTTGT 2534  
Qy SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTrpGlyValSerMetThrThr 780  
Db TCTGGCCCTAGACAGTCATTCAAATTTGGGTGGCATTTCTTATGGGGTTTCCATGAGACT 2594  
Qy LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTrpSer 800  
Db TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATATTAGT 2654  
Qy ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db GTATTTTACTCAGACGCCACTGGAGGAAGAAATGGATTCAAATCTTTTGTGAAATGGAA 2714  
Qy GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db CAGAGCTTACTTATCTTCTGAAACCAAGGCTKCTAGTAGAGCACCATGAGCCTTTGG 2774  
Qy LysGlnGlnMetFyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db AAACACACAGATGATACAAATAGCAAGTTTCTTCTTACCTTGAACGTAAGTAA 2834  
Qy SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db TCAGTGAGATCAGTGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2894  
Qy LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTrp 900  
Db TTGGTTTCACTCTCTTTTAAAAATGCTGTGGTTCCTCATCAAACTGTTTCCAGACTTATAT 2954  
Qy PheLeuLysProGlyAspLysProHisLysTrpLysThrSerLeuLeuLeuGlnAsnSer 920  
Db TTTCTAAAAACCTGGAGACAAACCAATAAATACAAAAACAAGTCTGCTTCTTCAAAATCT 3014  
Qy AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleValThr 940  
Db GCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAAGCCAGACATAATGTGTAGC 3074  
Qy MetIleAsnAspSerAspTrpValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db ATGATTTAATGACAGTGACTATGTATCCGTGGCTCCCATAGTGGCGCTTTAAATGTGTG 3134  
Qy HisSerGluLysAspTrpValPheAlaAlaValPheAsnSerThrMetValTrpSerLeu 980  
Db CATTCAGAAAAAGACTATGTTTTCAGCTGTTTTCACAAGCTACTATGTTTATTTCTTTA 3194  
Qy ProIleLeuValAsnIleIleSerAsnTrpTrpLeuTrpHisLeuAsnValThrGlnThr 1000  
Db CCTATATTAGTGAATATCAATTAGTAACACTACTATCTTTATCATTTAAATGTGACTGAACC 3254  
Qy IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db ATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTAATGATATAGTTTAAAAATTTGAG 3314  
Qy LeuTrpPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTrpPheAla 1040  
Db CTGTATTTTCAAGCAGCTTTGCTTGAATCATTTGTACTGCAATCTTACTGCCACCTTACTTGC 3374

Db 4141 CAGATAAACCCCTTTGGCCAGATACATACATTGCGAGAACATATTTTGAAATTTATGAGCT 4200  
QY 1401 VallysGlyMetSerAlaSerAspMetlysluValleSerArgilleThrHisAlaLeu 1420  
Db 4201 GTCAAGGAATGAGTGAAGTGAATGAAGAAGTCAATAGTGAATGAATGAATGACATCT 4260  
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlylleLysArgLys 1440  
Db 4261 GATTATAAAGAACATCTTCAGAGACTGTAAAGAACTACCTGCGAGGAATCAACAGAAAG 4320  
QY 1441 LeuCysPheAlaLeuSerMetLeuGlnProGlnLysThrLeuLeuAspGluProSer 1460  
Db 4321 TTGTGTTTGTCTAAGTATCTAGGAATCCTCAGATTACTTTGCTAGATGAACCATCT 4380  
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaLysThrAlaPheLys 1480  
Db 4381 ACAGTATGATCCCAAGCCCAACAGCACATGTGGGAGCAATTCGAATCGCATTTAAA 4440  
QY 1481 AsnArgLysArgAlaAlaLysLeuThrHisThrMetGluGluAlaGluAlaValCys 1500  
Db 4441 AACAGAAAGCGGCTGCTATTCTGACCACTCACTATATGAGAGGAGGAGGCTGTCTGT 4500  
QY 1501 AspArgValAlaLysMetValSerGlyLysLeuArgCyslleGlyThrValGlnHisLeu 1520  
Db 4501 GATCAGTAGCTATCATGTGTCTGGGAGTAAAGATGTATCGAACAGTACACATCTA 4560  
QY 1521 LysSerLysPheGlyLysGlyThrPheLeuGlnLysLysLeuLysAspTrpIleGluAsn 1540  
Db 4561 AAGAGTAAATTTGAAAAGGCTACTTTTGGAAATTAATTTGAAGAGCTGGATAGAAAC 4620  
QY 1541 LeuGluValAspArgLeuGlnArgGluLysGlnThrPheProAsnAlaSerArgGln 1560  
Db 4621 CTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCGCTCAG 4680  
QY 1561 GluSerPheSerSerIleLeuAlaTrpLysIleProLysGluAspValGlnSerLeuSer 1580  
Db 4681 GAAAGTATTTCTCTATTTGGCTTATATAAATTTCTAAGGAAGATGTTTCAGTCCCTTCA 4740  
QY 1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaLysGluLysThrSerPhe 1600  
Db 4741 CAATCTTTTAACTCGAAGAGCTAAACATGCTTTTGCATTTGAAGATATAGCTTT 4800  
QY 1601 SerGlnAlaThrLeuGlnValPheValGluLeuThrLysGluGlnGluAsp 1620  
Db 4801 TCTCAAGCAACATTCGAACAGGTTTTTGTAGAACTCACTAAAGAACCAAGAGGAGAGAT 4860  
QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640  
Db 4861 AATAGTTTGGAACTTTAAACAGCACACTTTTGGTGGGAACCAACAGAGATAGAGTA 4920  
QY 1641 ValPhe 1642  
Db 4921 GTATTT 4926

## RESULT 6

US-09-971-121-5  
; Sequence 5, Application US/09971121  
; GENERAL INFORMATION:  
; APPLICANT: Hu, Yi  
; APPLICANT: Nepomnichy, Boris  
; TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding the  
; TITLE OF INVENTION: Same  
; FILE REFERENCE: LEX-0250-USA  
; CURRENT APPLICATION NUMBER: US/09/971,121  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/239,629  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 5262  
; TYPE: DNA  
; ORGANISM: homo sapiens

## US-09-971-121-5

## Alignment Scores:

Prod. No.: 0 Length: 5262  
Score: 8395.00 Matches: 1637  
Percent Similarity: 99.70% Conservative: 0  
Best Local Similarity: 99.70% Mismatches: 5  
Query Match: 99.63% Indels: 0  
DB: 41 Gaps: 0

## US-10-090-458-5 (1-1642) x US-09-971-121-5 (1-5262)

QY 1 MetSerThrAlaAlaLeuArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
Db 255 ATGTCCACTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACCAGAACACTTCTACTGAG 314  
QY 21 AsnThrLeuLysCysArgThrLysLysSerValGlnGluLysLeuPheProLeu 40  
Db 315 AATTACTTAATTAATGAGAACCAAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 374  
QY 41 PhePheLeuPheTrpLeuLysLeuLysMetMetHisProAsnLysLysTrpGluGlu 60  
Db 375 TTTTCTTTTATTTGGTTTATATTAATAGCATGTGATGCCAATTAAGAAATATGAGAA 434  
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLysLeuGly 80  
Db 435 GTGCTTAATATAGAACTCAATCTATGGCAAGTTTACTCTTTCTAATCTAATCTTGA 494  
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
Db 495 TATACTCCAGTGACTAATATTAACAGCAGCATCATGCAAGAGTGTCTACTGATCATCTA 554  
QY 101 ProAspValIleThrGluGluTrpThrAsnGluLysGluMetLeuThrSerSerLeu 120  
Db 555 CTTGATGTCATATTTACTGAGAGATATACAAATGAAAGAAATGTTAACTCCAGTCTC 614  
QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTrpGluLeuArg 140  
Db 615 TCTAAGCCGAGCAACTTTCTAGGTGTGGTTTTTCAAGAGCTCCATCTCTATGAATCTGT 674  
QY 141 PhePheProAspMetIleProValSerSerIleTrpMetAspSerArgAlaGlyCysSer 160  
Db 675 TTTTCTTCCGTATGATTTCCAGTATCTTCTATTTATATGATTTCAAGAGCTGGCTGTCA 734  
QY 161 LysSerCysGluAlaAlaGlnTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
Db 735 AAATCATGTGAGGCTGCTCTCAGTACTGCTCCTCAGGTTTCACAGTTTACACATCCATA 794  
QY 181 AspAlaAlaIleThrGlnLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
Db 795 GATGCTGCCATATATACAGTTGAAGCAATGTTCTTTTGGAGAGAGCTGGAGTCAACT 854  
QY 201 LysAlaValIleMetGlyGluThrAlaValGluIleAspThrPheProArgGlyVal 220  
Db 855 AAAGCTGTTTATATGGAGAACTGCTGTTGTAGAAATAGATACCTTTTCCCGAGGAGTA 914  
QY 221 IleLeuIleTrpLeuValleAlaPheSerPropheGlyThrPheLeuAlaIleHisIle 240  
Db 915 ATTTTAAATATACCTAGTTATAGCATTTTCCACCTTTTGGTACTTTTGGCAATTCATATC 974  
QY 241 ValAlaGluLysGluLysLysIleLysGluPheLysIleValMetGlyLeuHisAspThr 260  
Db 975 GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATATATGGGACTTCAATGATCT 1034  
QY 261 AlaPheTrpLeuSerTrpValLeuLeuTrpThrSerLeuIlePheLeuMetSerLeuLeu 280  
Db 1035 GCCTTTTGGCTTTCTCTGGGTCTTCTATAWACAAGTTTAAATTTTCTTATGCTCCCTCT 1094  
QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
Db 1095 ATGGCAGTCAATGGCAGCTCTTTTGTATTATTTCTCAAGTAGCAGCATTTGTGATTT 1154  
QY 301 LeuLeuPhePheLeuTrpGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320

QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
DB 1981 AAAGCCAAATCGGTCGACAGTGTTCACTACTCAATTCATGGATGAAGCTGACATCTTGCA 2040  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
DB 2041 GATAGGAAGCTGTGATATCAAGAAGATGCTGAAGATGTGTGGTCTCTCAATGTCTCTC 2100  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 2101 AAAAGTAATAGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTTGGCCACA 2160  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
DB 2161 GAATCTCTTCTCTCACTGGTTAAACAACATATACCTGGAGTACTTTTATTAACAAGAA 2220  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
DB 2221 GACCAACAACCTTGTGTATAGCTTGTCTTCAAGGACATGGCAAAATTTTCAGGTTTGTT 2280  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 2281 TCTGCCCTAGACAGTCATTCAAATTTGGTGGCATTTCTTATGGGGTTTCCATGACGACT 2340  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
DB 2341 TTGGAGAGCTATTTTAAAGCTAGAGTTGAAGCAGAAATTTGACCAAGCAGATTTATAGT 2400  
QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 2401 GTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGGATTCAAAATCTTTTGTATCAATGGAA 2460  
QY 821 GlnSerLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
DB 2461 CAGAGCTTACTTATCTTCTGAAACCAAGGCTKCTCTAGTGAGCACCATGAGCCTTTGG 2520  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
DB 2521 AAACAACAGATGATACATAGCAAGATTTCAATTTCTTTACCTTGAAACGCTGAAAGTAAA 2580  
QY 861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 2581 TCAGTGAGATCAGTGTTGCTTCTGCTTTTAATTTTTTTTTCACAGTTTCAGATTTTATGTTT 2640  
QY 881 LeuValHisSerPheLysAsnAlaValAlaValProIleLysLeuValProAspLeuTyr 900  
DB 2641 TTGGTTTCATCACTCTTTTAAATAATGCTGGTTCCCATCAAACTTGTTCAGACTTATAT 2700  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
DB 2701 TTTCCTAAAACCTGGAGACAAACCAATAAATCAAAACAAAGTCTGCTTCTTCAAAATTC 2760  
QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
DB 2761 GCTGACTCAGATATCAGTCATCTTATAGCTTTTTCACAAGCCAGAACATATATGTTGACG 2820  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
DB 2821 ATGATTAATAGACAGTACATGATATCCGTTGGTCCCAATAGTGGCCTTTAAATGTGTG 2880  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
DB 2881 CATTCAGAAAAGACTATGTTTTTGAGCTGTTTTTCAACAGTACTATGTTTATTTCTTTA 2940  
QY 981 ProIleLeuValAsnIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
DB 2941 CCTATATAGTAATATCATATAGTAATCTATCTTTATCATTTAAATGTGACTGAAC 3000  
QY 1001 IleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
DB 3001 ATCCAGATCTGGAGTACCCCACTCTTTCAGAAATTTACTGATATAGTTTTTAAATGAG 3060  
QY 1021 LeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAla 1040

DB 3061 CTGTATTTTCAAGCAGCTTGTCTTGGAAATCATTTGTACTGCAATGCCACCTTACTTTGCC 3120  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
DB 3121 ATGGAAAATGCAGAGAATCATTAAGATCAAAAGCTTAYACTCAACTTAAACTTTCAGGCTTT 3180  
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleIle 1080  
DB 3181 TTGCCATCTGCATATTTGGATTTGGCAAGCTGTGTGTATATCCCTTATTTTATTCATT 3240  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
DB 3241 CTTATTTTGTCTAGGAAGCTTATGGCAATTCATTATGATGATTTATTTTATACTGTA 3300  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
DB 3301 AAGTTCTCTGTCTGTGTTTTCCTTATTTGGTATTTGTCATCAGTATTTCTGTTCATT 3360  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
DB 3361 TATATTGCTTCTTTCACCTTTAAGAAAATTTTAAATACCAAGAAATTTTGGTCAATTTATC 3420  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
DB 3421 TATTTCTGTGCGACGCTGGCTTGTATTTGCAATCAGTCAAAATAACTTCTTTTATGGATAC 3480  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
DB 3481 ACAATTGCAACTATTTCTTCAATATGCTTTTGTATCATCATTCATTCATTCATCTCTA 3540  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
DB 3541 GGTTCGCTGATTTCTTTTCAAGAAATTTCTTGGAGAATTTGAGGAGCAGATCAATAAGAAAAGAT 3600  
QY 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
DB 3601 TATAATCCATGGATAGGCTTTTCAGTAGCTGTATATCGCTTACCTGAGTGTGTACTG 3660  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
DB 3661 TCGATTTTCTCTTACAATACTATGAGAAAAAATATGGAGGAGCAGATCAATAAGAAAAGAT 3720  
QY 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAsp 1260  
DB 3721 CCCTTTTTCAGAAACCTTTTCAGCAAGCTTAAAAATAGGAAGCTTCCAGAACCCAGAC 3780  
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
DB 3781 AATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTTAAAGCTCAAAGAGCTGATGGT 3840  
QY 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
DB 3841 TGCACGTGTTTGTAGAGAGAAACCATTCATATGCTGAGCAATTTGCAATAAGAAATGAT 3900  
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysValLysValAlaThrLysTyrIleSer 1320  
DB 3901 GACAGAAAGATTTTCTTCTTTTCAAGAAAAGTAAAGAAAAGTGGCAACTTAAATACATCTCT 3960  
QY 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
DB 3961 TTCTGTGTGAAAAAGAGAGATCTTAGGACTATTTGGTCCCAATTTGGTGTGTGTCGCAAAAGC 4020  
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
DB 4021 ACAATTAATTAATTTCTGGTTGTGTATATTGAACCAACTTCAGGCCAGGATTTTTTAGGA 4080  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
DB 4081 GATTATTTCTTCAGAGACAGTGAAGATGATGATTCATCACTGAAGTGTATGGTGTACTGCTCT 4140  
QY 1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400



```

; TYPE: DNA
; ORGANISM: homo sapiens
US-09-971-121-1

Alignment Scores:
  Pred No.:      0      Length:      4929
  Score:      8395.00      Matches:      1637
  Percent Similarity:      99.70%      Conservative:      0
  Best Local Similarity:      99.70%      Mismatches:      5
  Query Match:      41      Indels:      0
  DB:      41      Gaps:      0

US-10-090-458-5 (1-1642) x US-09-971-121-1 (1-4929)

QY      1      MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20
DB      1      ATGTCCTCACTGCAATAGGAGGTAGGAGTTGGAGACAGACCAGAACACACTTCTACTGAAG 60

QY      21      AenTYrLeuIleLeuLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu 40
DB      61      AATTACTTAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAAATTCCTTTCCACTA 120

QY      41      PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysValGluGlu 60
DB      121      TTTTCTTTTATTTGGTTAATTAATAGCATGATGCATCCAAATAGAAATATGAAGAA 180

QY      61      ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80
DB      181      GTGCTCTAATATAGAACTCAATCCTATGACAAAGTTTACTCTTTCTAATCTAATCTTTGGA 240

QY      81      TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
DB      241      TATACTCCAGTGACTAATATTACAGCAGCATCATGACAGAAAGTGTCTACTGATCATCTA 300

QY      101      ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
DB      301      CCTGATGTCATANTTACTGGAAGATATACAAATGAAAAGAAATGTTAAACATCCAGTCTC 360

QY      121      SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
DB      361      TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAAAGACTCCATGTCCTATGAACCTTCGT 420

QY      141      PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
DB      421      TTTTCTCTGATATGATATCCAGTATCTTCTATTATATATGATTCAGAGCTGGCTGTTC 480

QY      161      LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180
DB      481      AAATCATGTAGGCTGCTCAGTACTGTGCTCCTCAGGTTTTCACAGTTTTCACAAAGCATCCATA 540

QY      181      AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200
DB      541      GATGCTGCCATATATACAGTTGAAGACCAATGTTCTCTTTGAGAGGAGCTGGAGTCAACT 600

QY      201      LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
DB      601      AAAGCTGTATTATGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGGAGGAGTA 660

QY      221      IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
DB      661      ATTTTAATATACCTAGTTATAGCATTTTCCACTTTTGGATATCTTTTGGCAATTCATATC 720

QY      241      ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
DB      721      GTAGCAGAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 780

QY      261      AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280
DB      781      GCTTTTGGCTTCTCTGGGTTCTTCTATAWACAAGTTTAAATTTTCTTATGCTCCCTCTT 840

QY      281      MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300
DB      841      ATGGCAGTCATTCGACACACTTCTTTGTATTCTCTCAAGTAGCAGCATTTGTGATATT 900

```

```

QY      301      LeuLeuPhePheIleuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320
DB      901      CTGCTTTTTCCTTTATGGATTATCACTGATTTTGTCTTTATGCTGACACTCTT 960

QY      321      PheLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340
DB      961      TTTAAAAAATCAAAACATGCTGGGAATAGTTGAATTTTGTACTGTGGCTTTTGATTT 1020

QY      341      IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360
DB      1021      ATTGGCTTATGATATCTCATAGAAAGTTTCCAAATCGTTTGTAGTGGCTTTTCAGT 1080

QY      361      ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380
DB      1081      CCTTCTGTCACCTGACTTTTGTGATGTTGTTGACAGGTATGCTTATAGAGATTTT 1140

QY      381      AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400
DB      1141      AATGAAGGTGCTTCATTTTCAAATTTGACTGAGGCCCATATCCCTCAATATTACAAAT 1200

QY      401      IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTrpLeuAspGlnVal 420
DB      1201      ATCATGCTCACACTTAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTGATCAAGTC 1260

QY      421      IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440
DB      1261      ATTCCAGGGGAAATTTGGCTTACGGAGATCATCTTATATATTTTCTGAAGCTTCATATTGG 1320

QY      441      SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460
DB      1321      TCNAAGAGCAAAAGAAATTTATGAGNGTTATCAGAGGCAATGTTAATGGAATATTAGT 1380

QY      461      PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480
DB      1381      TTTAGTGAATATTATCAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAACCCATAAGAAT 1440

QY      481      SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500
DB      1441      AGTGGTATTTCAGAAAGACATACAGAAAGAGGTGAATGTGGAGGCTTTTGAGAAATTTG 1500

QY      501      SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520
DB      1501      TCATTTGACATATATGAGGTCAGATTACTGCTGCTTACTTGGCCACAGTGGAAACAGGAAAG 1560

QY      521      SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540
DB      1561      AGTACATTTGATGAATATCTTTGTGGACTCTGCCACCTTCTGTGGGTTTGTCACTATA 1620

QY      541      TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560
DB      1621      TATGGACACAGAGTCTTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATTTGCAATT 1680

QY      561      CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580
DB      1681      TGTCACAGTTAGATATACACTTTGATGTTTTCACAGTAGAAGAAATTTATCAATTTTG 1740

QY      581      AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600
DB      1741      GCTTCAATCAAGGGATACACGCCCAATATAATACAGAGAGTGCAGAGGTTTACTATA 1800

QY      601      AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys 620
DB      1801      GATTTAGACATGACAGTATCAAGATAACCAAGCTTAAAAAATTAAGTGGTGGTCAAAA 1860

QY      621      ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640
DB      1861      AGAAAGCTGTCTATTAGGAATGCTGTTCTTGGAGCCCAAGATATCTGCTGTAGTAGAA 1920

QY      641      ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660
DB      1921      CCAACAGCTGAATGGACCCCTGTTCTCGACATATTGTATGGAATCTTTTAAAAATACAGA 1980

```

QY 1021 LeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 4071 CTGTATTTTCAGACGCTTTGCTTGAATCATTTGTACTGCAATGCCACCTTACTTTGCC 4130  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 4131 ATGGAATATGCAGAGAAATCAATGAATCAAGCTTATCTCAACTTTAAACTTTTCAGTCTT 4190  
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleIle 1080  
Db 4191 TTGCCATCTGCATATTGGATTGGACAGCTGTGTGTGATATCCCTTATTTTATATCAT 4250  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 4251 CTATTTTTCATGCTAGCAAGCTTACTGGCATTTTCATTAATGAGATTATATTTTATATCTGA 4310  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 4311 AAGTTCCTTGT 4370  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 4371 TATATTGCTTCTTACCTTTAAGAAATTTTAAATACCAAGAAATTTTGTTCATTTATC 4430  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 4431 TATTTCTGTGGCGCTTGT 4490  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
Db 4491 ACAATTTGCACATATCTTCAATATGCTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4550  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
Db 4551 GGTGTGCTGATTTCTTTTCATAAGATTTCTTGAAGAATGTAGAGGAGCATCAATAGAAAGAT 4610  
QY 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
Db 4611 TATAATCCATGGGATAGGCTTTTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4670  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
Db 4671 TGGATTTTCTCTTCAATACTATGAGAAATAATATGAGGAGCATCAATAGAAAGAT 4730  
QY 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
Db 4731 CCTTTTTCAGAAACCTTTTCAGGAAGTCTTAAATATAGGAAGCTTCCAGAACCCAGAC 4790  
QY 1261 AsnGluAspGluAspGluValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 4791 AATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGACTAAAGGTCAAAGAGCTGATGGT 4850  
QY 1281 CysGlnCysCysGluCylulysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
Db 4851 TGCACGTGTGTGAGGAGAAACCATCCATATGTCAGCAATTTGCATAAAGAATATGAT 4910  
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer 1320  
Db 4911 GACAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCAACTAAATATACATCTCT 4970  
QY 1321 PheCysValLysLysGlyCylulleuGlyLeuLeuGlyProAsnGlyAlaCylLysSer 1340  
Db 4971 TTTCTGTGAAAAAGAGAGATCTTAGACATATTGGGTCCAAATGGGTGGGCAAAAGC 5030  
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 5031 ACAATTTATATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5090  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
Db 5091 GATTATTTCTTCAGAGCAAGTGAAGATGATGATTCATGAGTGTATGGTTACTGTCTCT 5150

QY 1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
Db 5151 CAGATAAACCTTTTGTGGCAGATACATACATTCGAGAAACATTTTGAATTTATGGAGCT 5210  
QY 1401 ValIysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db 5211 GTCAAGAGATAGTCAAGTGCATGAAGAAGTCAATGAAGTCAATGAAGTCAATGAAGTCA 5270  
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
Db 5271 GATTTAAAGAAACATCTTCAGAGACTGTAAAGAACTACCTGCAGGAATCAACCAAGAG 5330  
QY 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
Db 5331 TTGTGTTTCT 5390  
QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
Db 5391 ACAGGTATGGATCCCAAGCCAAACAGCACATGTGGGAGCAATTCGAACTGCATTTTAA 5450  
QY 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500  
Db 5451 AACAGAAAGCGGCTGCTATTTCTGACCACTCATATATGAGGAGGAGGAGGCTGTCTGT 5510  
QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
Db 5511 GATCGAGTAGCTATCATGTGTCTGGCAGTTAAGATGTATCGAACAGTACAACTCTA 5570  
QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
Db 5571 AAGAGTAAATTTTGAAGAGCTACTTTTGGAAATTTAAATGAAGCACTGGATAGAAAC 5630  
QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlyTyrIlePheProAsnAlaSerArgGln 1560  
Db 5631 CTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATCAAGCCCTCAG 5690  
QY 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
Db 5691 GAAAGTTTCTTCT 5750  
QY 1581 GluSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluLysSerPhe 1600  
Db 5751 CAATCTTTTAAAGTGGAGAAAGCTAAACATGCTTTTGCCATTTGAAGATATAGCTTT 5810  
QY 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
Db 5811 TCTCAGCAACATTTGAACAGGTTTGTAGAACTCCTTAAGAACACAGGAGGAGAGAT 5870  
QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640  
Db 5871 AATAGTTGTGGAACCTTTAAACAGCACCTTTGTGGGAAACGAACACAGAAAGATAGATA 5930  
QY 1641 ValPhe 1642  
Db 5931 GTATTT 5936

RESULT 5  
US-09-971-121-1  
; Sequence 1, Application US/09971121  
; GENERAL INFORMATION:  
; APPLICANT: Hu, Yi  
; APPLICANT: Nepomichy, Boris  
; TITLE OF INVENTION: Novel Human Transporter Proteins and Polynucleotides Encoding th  
; FILE REFERENCE: Same  
; FILE REFERENCE: LEX-0250-USA  
; CURRENT APPLICATION NUMBER: US/09/971,121  
; CURRENT FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: US 60/239,629  
; PRIOR FILING DATE: 2000-10-10  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 4929

1851 ATGGGAGTCATTGGCAGACAGCTCTTTGTTATTTCTCAAGTAGCAGCAATTTGTGATATTT 1910  
301 LeuLeuPhePheLeuTyrglyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
1911 CTGCTTTTTCCTTTATGGATATCATCTGTATTTTGTCTTTAATGCTGACACCTCTT 1970  
321 PheLysLysSerLysHisValGlyLeValGluPhePheValThrValAlaPheGlyPhe 340  
1971 TTTAAAAAATCAAAACATGTGGGATAGTGTGAATTTTTTGTCTGTGGCTTTTGGATTT 2030  
341 lleGlyLeuMetillelleLeulleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
2031 ATTGGCTTATGATTAATCTCATGAAGAAGTTTCCCAAAATCGTTAGTGTGGCTTTTCAGT 2090  
361 ProPheCysHisCysThrPheValilleGlylleAlaGlnValMetHisLeuGluAspPhe 380  
2091 CTTTCTGTCACTGTACTTTTGTGATTTGGTATGTGCACAGGTCATGCATTTAGAAGATTTT 2150  
381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProfyProLeuillelleThrile 400  
2151 AATGAAGGTGCTTCAATTTCAAAATTGACTGCAGGCCCATATCTCTAAATATTACAAT 2210  
401 lleMetLeuThrLeuAsnSerillePheTyrrValLeuLeuAlaValTyrrLeuAspGlnVal 420  
2211 ATCATGCTCACACTTAATAGTATATTCATGTCTCTTGGCTGTCTATCTTGTATCAAGTC 2270  
421 lleProGlyGluPheGlyLeuArgSerSerLeuTyrrPheLeuLysProSerTyrrP 440  
2271 ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCCTTCATATTGG 2330  
441 SerLysSerLysArgAsnTyrrGluLeuSerGluGlyAsnValAsnGlyAsnLysSer 460  
2331 TCAAAAGATTAAGAAATATGAGAGATTATGAGAGTTATCAGAGGCAATGTTAATGGAAATATTAGT 2390  
461 PheSerGluillelleGluProValSerSerGluPheValGlyLysGluAlaAilArgile 480  
2391 TTTAGTGAATTTATTGAGCCAGTTCTTTCAGATTTTGTAGGAAGAGCCATTAAGATTT 2450  
481 SerGlylleGlnLysThrTyrrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
2451 AGTGGTATTTCAAGAAGACATACAGAAGAAGGGTGAAAATGTGGAGGCTTTGAGAAATTTG 2510  
501 SerPheAspilleTyrrGluGlyGlnilleThrAlaLeuLeuGlyHisSerGlyTyrrGlyLys 520  
2511 TCATTTGACATATATGAGGGTCAGATTACTGCTTACTTTGGCCACAGTGGACAGGAAG 2570  
521 SerThrLeuMetAsnilleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerile 540  
2571 AGTACATTTGATGAATATTTCTTTGTGGACTCTGCCACCTTCTGTATGGGTTTGCATCTATA 2630  
541 TyrrGlyHisArgValSerGluilleAspGluMetPheGluAlaArgLysMetilleGlylle 560  
2631 TATGGACACAGAGTCTCAAAATAGATGAATGTTTGAAGCAAGAAATGATTTGGCAT 2690  
561 CysProGlnLeuAspilleHisPheAspValLeuThrValGluGluAsnLeuSerilleLeu 580  
2691 TGTCCACAGTTAGATATACATTTGATGTTTGCAGTAGAAGAAATTTATCAATTTTG 2750  
581 AlaSerilleLysGlylleProAlaAsnillelleGlnGluValGlnLysValLeuLeu 600  
2751 GCTTCAATCAAGGGATACCAAGCAACAATATAATACAAAGAGTGCAGAAAGTTTACTA 2810  
601 AspLeuAspMetGlnThrilleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
2811 GATTTAGACATGCAGACTATCAAGATTAACAAGCTAAAAAATTAAGTGTGTGTCAAAAA 2870  
621 ArgLysLeuSerLeuGlylleAlaValLeuGlyAsnProLysilleLeuLeuLeuAspGlu 640  
2871 AGAAGCTGTCTATTAGGAATTTGCTTCTTGGGAACCCAAAGATACGTGCTGTAGATGA 2930  
641 ProThrAlaGlyMetAspProCysSerArgHisilleValTrpAsnLeuLeuLysTyrrArg 660

2931 CCAACAGCTGGAATGACCCCTGTTCTTCGACATATTTGTATGGAATCTTTTAAAAATACAGA 2990  
661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspLysLeuAla 680  
2991 AAAGCCAAATGGGTGACAGTGTTCAGTACTCTATTTTCATGGATGAAGCTGACATTTCTTGA 3050  
681 AspArgLysAlaValilleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
3051 GATAGGAAGCTGTGATATCACAGGAATGCTGAAATGTGTGTTCTTTCAATGTTCCTC 3110  
701 LysSerLysTrpGlylleGlyTyrrArgLeuSerMetTyrrilleAspLysTyrrCysAlaThr 720  
3111 AAAAGTAAATGGGGATCGGCTACCGCTCGAGCATGTACATAGACAAATATTGTGTCACA 3170  
721 GluSerLeuSerSerLeuValLysGlnHisilleProGlyAlaThrLeuLeuGlnAsn 740  
3171 GAATCTCTTCTCATCTGGTTAAACAACATATATACCTGGAGCTACTTTATTACACAGAT 3230  
741 AspGlnGlnLeuValTyrrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
3231 GACCAACAATTTGTGTATAGTTCCTTTCAAGGACATGGACAAATTTTCAGGTTTGT 3290  
761 SerAlaLeuAspSerHisSerAsnLeuGlyValilleSerTyrrGlyValSerMetThrThr 780  
3291 TCTGCCCTAGACAGTCAATTCAAATTTGGGTGTCAATTTCTATGGTGTTCCTATGAGCACT 3350  
781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluilleAspGlnAlaAspTyrrSer 800  
3351 TTGGAAGACGTATTTTAAAGCTAGAAGTTGAAGCAGAAATTCACCAAGCAGATATAGT 3410  
801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
3411 GTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTGTATGAATGGAA 3470  
821 GlnSerLeuLeuilleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
3471 CAGAGCTTACTTATTTCTTCTGAACCAAGGCTTCTCTAGTCAGCACCATGAGCCTTGG 3530  
841 LysGlnGlnMetTyrrThrilleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
3531 AAACAACAGATGTATACAAATAGCAAGTTTCATTTCTTTACCTTGAACCTGAAAGTAA 3590  
861 SerValArgSerValLeuLeuLeuLeuLeuPhePheThrValGlnllePheMetPhe 880  
3591 TCAGTCAGATCAGTGTGCTTCTGCTTTTAATTTTTCACAGTTTCAGATTTTATGTT 3650  
881 LeuValHisSerPheLysAsnAlaValValProilleLysLeuValProAspLeuTyrr 900  
3651 TTGGTTTCATCACTCTTTTAAAAATGCTGTGGTTCCTCATCAAACTTGTTCACAGCTTATAT 3710  
901 PheLeuLysProGlyAspLysProHisLysTyrrLysThrSerLeuLeuLeuGlnAsnSer 920  
3711 TTTCTAAAAACCTGGAGAACCAACCAATAAATACAAAACAAAGTCTGCTTCTTCAAAATCT 3770  
921 AlaAspSerAspLysSerAspLeulleSerPhePheThrSerGlnAsnilleMetValThr 940  
3771 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAAAGCCAGAACATAATGCTGAGC 3830  
941 MetilleAsnAspSerAspTyrrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
3831 ATGATTAATGACAGTACTATGTCGGTGGCTCCCATAGTCGGCTTTTAAATGTGATG 3890  
961 HisSerGluLysAspTyrrValPheAlaAlaValPheAsnSerThrMetValTyrrSerLeu 980  
3891 CATTCAGAAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTATGTTTATTTCTTTA 3950  
981 ProilleLeuValAsnilleSerAsnTyrrLeuTyrrHisLeuAsnValThrGluThr 1000  
3951 CTTATATTAGTGAATATCATTAAGTAACTACTATCTTTATCATTTAAATGTGACTGAAACC 4010  
1001 lleGlnilleTrpSerThrProPhePheGlnGluilleThrAspilleValPheLysilleGlu 1020  
4011 ATCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTAATGATATAGTTTTPAAAAATTGAG 4070

```
Db 5498 GAITTTAAAGAACATCTTCAGAGACTGTAAAGAACTACTCGGAGATCAACAGCAAG 5557
Qy 1441 leuCysPheAlaLeuSerMetLeuGlyAenProGlnIleThrLeuLeuAseGluProSer 1460
Db 5558 TTGTCTTTTGTCTCTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCATCT 5617
Qy 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTpArgAlaIleArgThrAlaPheLys 1480
Db 5618 ACAGGTATGGATCCCAAGGCCAAACAGACATGTGGCGAGCAATTCGAACTGCAATTAAA 5677
Qy 1481 AenArgLysArgAlaAlaIleLeuThrThrHisTyMetGluGluAlaGluAlaValCys 1500
Db 5678 AACAGAAAGCGGCGTCTATTCTGACCCTCACTATATGGAGGAGCGAGCGCTGTCTGT 5737
Qy 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520
Db 5738 GATCGGTAGCTATCATGGTGTCTGGCAGTTAGATGTTATCGGACAGTACACACTTA 5797
Qy 1521 LysSerLysPheGlyLysGlyTyThrPheLeuGluLysLeuLysAspTrrPileGluAse 1540
Db 5798 AAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTTAAATTTGAAGGACTGGATAGAAAAC 5857
Qy 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyThrPheProAseAlaSerArgGln 1560
Db 5858 CTAGAAGTAGCCGCTTCAAGAGAAATTCAGTATATTTCCCAATTCGAAAGCCGTCAG 5917
Qy 1561 GluSerPheSerSerIleLeuAlaTyrlsleProLysGluAseValGlnSerLeuSer 1580
Db 5918 GAAAGTATTTCTCTATTTTGGCTATATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTCA 5977
Qy 1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTySerPhe 1600
Db 5978 CAATCTTTTTPAAGCTGAAGAGCTAAACATGCTTTTGCCATTTGAAGATATAGCTTT 6037
Qy 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAse 1620
Db 6038 TCTCAAGCAACATTCGACAGCTTTTGTAGAACTCACTAAAGAACAAAGAGGAGGAGAT 6097
Qy 1621 AenSerCysGlyThrLeuAseSerThrLeuTrpGluArgThrGlnGluAseArgVal 1640
Db 6098 AATAGTTGTGGAACCTTAAACAGCACACTTTGGTGGGAACCAACACAGAAAGATAGAGTA 6157
Qy 1641 ValPhe 1642
Db 6158 GTATT 6163
RESULT 4
US-10-005-338B-1
; Sequence 1, Application US/10005338B
; GENERAL INFORMATION:
; APPLICANT: DENEPELE, Patrice
; APPLICANT: ROSTER-MONTUS, Marie-Francoise
; APPLICANT: PRADES, Catherine
; APPLICANT: ARNOULD-REGUIGNE, Isabelle
; APPLICANT: DUVERGER, Nicolas
; APPLICANT: ALLIKMETS, Rando
; APPLICANT: DEAN, Michael
; TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENES
; TITLE OF INVENTION: CONTAINING SUCH NUCLEIC ACIDS, AND USES THEREOF
; FILE REFERENCE: ABCA5 6, 9, 10
; CURRENT APPLICATION NUMBER: US/10/005,338B
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/263,231
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: FR 00403440.1
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6525
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 4449
; OTHER INFORMATION: n=unknown, may be a or s or c or t
US-10-005-338B-1
Alignment Scores:
Pred. No.: 0 Length: 6525
Score: 8422.00 Matches: 1641
Percent Similarity: 99.94% Conservative: 0
Best Local Similarity: 99.94% Mismatches: 1
Query Match: 99.95% Indels: 0
DB: 43 Gaps: 0
US-10-090-458-5 (1-1642) x US-10-005-338B-1 (1-6525)
Qy 1 MetSerThrAlaIleArgGluValGlyValTpArgGlnThrArgThrLeuLeuLys 20
Db 1011 ARGTCACATGCAATTTAGGAGGTAGAGTTTGGACACAGACCAACACTTCTACTGAAG 1070
Qy 21 AenTyrlsleLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu 40
Db 1071 AATTACTTAATTAATTCAGAACCAACCAAGAGTAGTGTTCAGGAAATCTTTTCCACTA 1130
Qy 41 PhePheLeuPheTrpLeuIleLeuSerMetMetHisProAseLysLysTyrlsGlu 60
Db 1131 TTTTTTTTATTTGGTTAATATTAATAGCATGATGCAATTCACCAATAAGAAATATGAAGAA 1190
Qy 61 ValProAseIleGluLeuAseProMetAspLysPheThrLeuSerAseLeuLeuGly 80
Db 1191 GTGCCTAATATAGAACTCAATCCTATGGACAAAGTTTACTCTTTCTAATCTAATCTTGA 1250
Qy 81 TyThrProValThrAseIleThrSerSerIleMetGlnLysValSerThrAseHisLeu 100
Db 1251 TATACTCCAGTGACTAATATTAACAGCAGCATCATCGAGAAAGTGTCTACTGATCATCTA 1310
Qy 101 ProAspValIleIleThrGluGluTyThrAseGlnLysGluMetLeuThrSerSerLeu 120
Db 1311 CCTGATGTCTAATTTACTCTGAAGAATATACAAATGAAAGAAATGTTAACATCCAGTCTC 1370
Qy 121 SerLysProSerAsePheValGlyValValPheLysAseSerMetSerTyrlsGluAse 140
Db 1371 TCTAAGCCGAGCAACTTTGTAGTGTGGTTTTCAGAAAGTCCCATGCTCTATGAATCTGT 1430
Qy 141 PhePheProAspMetIleProValSerSerIleTyMetAsePheArgAlaGlyCysSer 160
Db 1431 TTTTTTCTGTATATGATTCAGTATCTTCTATTTATATGGATTCAGAGCTGGCTGTCA 1490
Qy 161 LysSerCysGluAlaAlaGlnTyTrpSerSerGlyPheThrValLeuGlnAseIle 180
Db 1491 AAATCATGTGAGCGTCTCAGTACTGTCTCAGGTTTTCACAGTTTTCACAGCATCCATA 1550
Qy 181 AspAlaAlaIleIleGlnLysThrAseValSerLeuTrpLysGluLeuGluSerThr 200
Db 1551 GATGCTGCCATTTATACAGTTGAAGCAATGTTTCTCTTTGGAAGAGAGCTGGAGTCACT 1610
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAseThrPheProArgGlyVal 220
Db 1611 AAAGCTGTTATTATGGAGAAACTGCTGTGTGAGAAATAGATACCTTTCCCGAGGAGTA 1670
Qy 221 IleLeuIleTyrlsLeuValIleAlaPheSerProPheGlyTyThrPheLeuAlaIleHisIle 240
Db 1671 ATTTTAATATACCTAGTTATAGCATTTTCCCTTTTGGATACTTTTGGCAATTCATATC 1730
Qy 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
Db 1731 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATCATACT 1790
Qy 261 AlaPheTrpLeuSerTrpValLeuLeuTyThrSerLeuIlePheLeuMetSerLeuLeu 280
Db 1791 GCCTTTTGCTTCTCCTGGGTCTCTATATACAAAGTTTAAATTTTCTATGTCCTCTT 1850
Qy 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300
```

QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 3338 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATGTGCCACA 3397  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
DB 3398 GAACTCTCTTCTTCACCTGGTTAAACACATATACCTGGAGTACTTTATTACAACGAAT 3457  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
DB 3458 GACCAACAACCTGTGTATAGCTTCCCTTTCAGGACATGGCAAAATTTTCAGGTTTGTGT 3517  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 3518 TCTGCCCTAGACAGTCATTCAAAATTTGGTGTCTATTTCTTATGGTGTTCCTCAACGACT 3577  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
DB 3578 TTGGAAGACGTATTTTAAAGCTAGAAGTTGAACGAGAAATGTGACCAAGCAGATTATAGT 3637  
QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 3638 GTATTTACTCAGACGCCACTCGAGGAGAAATGGATTCAAATCTTTTGATGAATGGAA 3697  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
DB 3698 CAGAGCTTACTTATTTCTTCTGAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTGG 3757  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysAspGluSerLys 860  
DB 3758 AAACACAGATGATACANTAGCAAGTTTCATTTCTTACCTTGAACCGTGAAGTAAA 3817  
QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 3818 TCAGTGAGATCAGTGTCTCTGCTTTTAATTTTTTTTTCACAGTTCAGATTTTATGTTT 3877  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
DB 3878 TTGGTTTCATCACTCTTTTAAATAATGCTGTGTCTCCATCAAACTGTGTCAGACTTATAT 3937  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnGlnAsnSer 920  
DB 3938 TTTCTAAACCTGGAGACAAACCATATAATACAAACCAAGTCTGCTTCTTCAAAATCT 3997  
QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
DB 3998 GCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAAAGCCAGAACATAATGTCAGC 4057  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaLeuAsnValMet 960  
DB 4058 ATGATTATATGACATGACTATGTATCCCTGGCTCCCATAGTGGGCTTTAAATGTGGTG 4117  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
DB 4118 CATTCAGAAAAGGACTATGTTTTTGGCAGCTGTTTTCAACAGTACTATGTTTATCTTTA 4177  
QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
DB 4178 CCTATATTAGCAATATCATTAGTAACCTACTATCTTTATCATTTTAAATGTGACTGAAC 4237  
QY 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
DB 4238 ATCCAGATCTGGAGTACCCCTTCTTTCAGAAATTTACTGATATAGTTTTTAAAAATGAG 4297  
QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
DB 4298 CTGTATTTTCAAGCAGCTTTGCTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCC 4357  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
DB 4358 ATGGAATATGCAGAGAAATCATAAAGATCAAGCTTATATCTCAACTTAAACTTTTCAGGCTT 4417

QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleIle 1080  
DB 4418 TTGCCATCTGCATATTTGGATTGGACAGCTGTGTGTATATCCCTTATTTTATCATTT 4477  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
DB 4478 CTTATTTTGTATCTAGGAAGCTTATTTGGCATTTTCATTTATGGATTATATTTTATACTGA 4537  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
DB 4538 AGTTTCTTGTGTGGTTTTTTCCTTATTTGGTTATGTTCATCATCATTTCTGTCTACT 4597  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
DB 4598 TATATTGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCATTTATC 4657  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
DB 4658 TATCTGTGGCAGCGTTGGCTTGTATTTGCATCATCATGAAATAACTTTCTTTATGGATAC 4717  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
DB 4718 ACAATTTGCAACTTATTTCTTTCATTTATGCTTTTGTATCATCATTCATTCATCTCTA 4777  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTyrLysAsnValArgLysAsnValAspThr 1200  
DB 4778 GGTGTGCTGATTTCTTTCAATAAGATTTCTTGGAGAATGTACGAAATAATGTGACACC 4837  
QY 1201 TyrAsnProTyrAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
DB 4838 TATATCCATCGATAGGCTTTTCAGTAGCTGTATATGCTTACCTTACCTCATCTCTA 4897  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
DB 4898 TGGATTTTCTCTTACATATCTATGAGAAAAATATGAGGCGAGATCAATAAGAAAAAGAT 4957  
QY 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAsp 1260  
DB 4958 CCTTTTTCAGAAACCTTTCAACGAAGCTTAAAAATAGGAAGCTTCCAGAACCCACCAGAC 5017  
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
DB 5018 AATGAGATGAAGATGAAGATCTCAAGCTGAAGACTTAAAGGTCAAGAGCTGATGGT 5077  
QY 1281 CysGlnCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
DB 5078 TGGCAGTGTGTGAGGAGAAACCATTCATATGTCAGCATTTGCTAATAAGATATGAT 5137  
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320  
DB 5138 GACAAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAATACATCTCT 5197  
QY 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValagLysSer 1340  
DB 5198 TTTCTGTGCAAAAAAGGAGAGATCTTAGGACTATTGGGTCCAAATGGTGTGGCAAAAGC 5257  
QY 1341 ThrIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
DB 5258 ACAATTTATATTTCTGTTGCTGTATTTGAACCAACTTCCAGGCCAGGATTTTTTAGGA 5317  
QY 1361 AspTyrSerSerGluThrSerGluAspAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
DB 5318 GATTATTTCTTCAGAGCAAGTGAAGTATGATGATTCATGAAAGTGTATGGTTTACTGTCT 5377  
QY 1381 GlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
DB 5378 CAGATAAACCTTTTGGCCAGATACTACTATTCAGGAGAACATTTTGAATTTATGGAGCT 5437  
QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
DB 5438 GTCAAGGAATGAGTGAAGTGCATGAAAGTGAAGTCAATAAGTCAATAACATGCATCTT 5497  
QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440

Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.94%	Mismatches:	0
Query Match:	99.95%	Indels:	0
DB:	78	Gaps:	0
US-10-090-458-5 (1-1642) x US-60-223-269-15 (1-6369)			
Qy	1	MetSerThrAlaIleAysGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys	20
Db	1238	ATGTCCACTGCAATTAGGGAGTAGGAGTTGGAGACAGACCACACACATCTACTGAAG	1297
Qy	21	AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGluIleLeuPheProLeu	40
Db	1298	AATTACTTAAATAAATGAGAACCAAAAGAGTAGTGTTCCAGGAAATCTCTTTTCCACTA	1357
Qy	41	PhePheLeuPheTTPLeuIleLeuIleSerMetMethHisProAsnLysLysTyrGluGlu	60
Db	1358	TTTTTTTATTTGGTTAAATTAATTAGCATGATGTCATCAATAAGAAATAGAAGAA	1417
Qy	61	ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly	80
Db	1418	GTGCCTAATAATAGAACTCAATCCTATGGACAAGTTACTCTTTCTTAATCTTAATCTTGG	1477
Qy	81	TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu	100
Db	1478	TATACTCCAGTGACTAATATATACAGACGACATCATGCAGAAAGTGCTACTGATCATCTA	1537
Qy	101	ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu	120
Db	1538	CCTGATGTGCATTAATTTACTTGAAGAATATACAAATGAAAAGAAATGTTAAACATCCAGTCTC	1597
Qy	121	SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg	140
Db	1598	TCTAAGCCGAGCAACTTTGTAGGTGTGGTTTTCAAGACTCCATGTCCTATGAATTCGT	1657
Qy	141	PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer	160
Db	1658	TTTTTTCTCGATGATGATCCAGTATCTTCTATTTATATGAGTATCAAGAGCTGGGTGTTC	1717
Qy	161	LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle	180
Db	1718	AAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTACAGTTTACAGCATCCATA	1777
Qy	181	AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGluSerThr	200
Db	1778	GATGTCGCATTTATACAGTTGAAGACCAATGTCTCTTTTCGGAAGGAGCTGGAGTCAACT	1837
Qy	201	LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal	220
Db	1838	AAAGCTGTATTATGGGAGAACTGCTGTTGTAGBAATAGATACCTTTCCCGAGGAGTA	1897
Qy	221	IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle	240
Db	1898	ATTTTAAATATACCTAGTTATATAGCATTTTTCACCTTTTGATATCTTTTGGCAATTCATATC	1957
Qy	241	ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr	260
Db	1958	GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT	2017
Qy	261	AlaPheTyrLeuSerTyrValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu	280
Db	2018	GCCTTTTGGCTTTCTCTGGGTCTCTATATACAAAGTTTAAATTTTCTTATGTCCTCTCT	2077
Qy	281	MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe	300
Db	2078	ATGGCAGTCAATGGCAGCTCTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGTGATATT	2137
Qy	301	LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu	320
Db	2138	CTGCTTTTTTCCCTTATGGGAATATCACTGTATTTTTTGTCTTAAATGCTGACACCTCTT	2197
Qy	321	PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe	340

Db	2198	TTTTAAAAATCAAAACAATGTGGGAATAGTTGAATTTTTTTGTACTGTGGCTTTTGGATTT	2255
Qy	341	IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer	360
Db	2258	ATTGGCCTTTATGATAAATCCCTCATAGAAAGTTTTCCCAAAATCGTGTAGTGGCTTTTCAGT	2317
Qy	361	ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMethHisLeuGluAspPhe	380
Db	2318	CCTTCTGTCTACGTGACTTTTGTGATGTGATTTGCACAGGTCATGCATTTTAGAAGATTTT	2377
Qy	381	AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle	400
Db	2378	AATGAAGGTCCTTCATTTTCCAAATTTGACTGCAGGGCCATATCCTCTAAATATTAACAATT	2437
Qy	401	IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal	420
Db	2438	ATCATGCTCACACTTAATAGATATTTCTATGCTCTTGTGCTGTCTATCTTGATCAAGTC	2497
Qy	421	IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheIleuLysProSerTyrTrp	440
Db	2498	ATTCACGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTCTCGAAGCCTTCATATTTG	2557
Qy	441	SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer	460
Db	2558	TCAAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGGCAATGTTAATGAAATATTTAGT	2617
Qy	461	PheSerGluIleIleGluProValSerSerGluPheValGlyGluAlaIleArgIle	480
Db	2618	TTTASTGAAATATTGTAGCCAGTTTCTTCAGAAATTTGTAGAAAGAGGCCATAAGAAAT	2677
Qy	481	SerGlyIleGlnLysThrTyrArgLysGlyGluAsnValGluAlaLeuArgAsnLeu	500
Db	2678	AGTGATTCAGAAGACATACAGAAAGAGGGTGAAATCTCGAGGCTTTGAGAAATTTG	2737
Qy	501	SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys	520
Db	2738	TCATTTGACATATATGAGGGTCAGATTACTTGCCCTTACTTGGCCACAGTGAAACAGGAAG	2797
Qy	521	SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle	540
Db	2798	AGTACATTTGATGAATATCTTTGTGGACTCTGCCACCTTCTGATGGTTTGCATCTATA	2857
Qy	541	TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle	560
Db	2858	TATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAGCACAGAAATATGATTTGGCATT	2917
Qy	561	CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu	580
Db	2918	TGTCACAGTTAGATATACACTTTGATGTTTTGACACAGAGAAATTTATCAATTTTG	2977
Qy	581	AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu	600
Db	2978	GCTTCAATCAAAAGGGATACCAGGCCAAATATATAACAAGAAGTCAGAAAGGTTTACTA	3037
Qy	601	AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys	620
Db	3038	GATTTAGACATGCAGACTATCAAGAATAACCAAGCTAAAAAATTAAGTGTGTGTCACAAA	3097
Qy	621	ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuAspGlu	640
Db	3098	AGAAAGCTGTCTATTAGGAATTTGCTGTCTTGGGAACCCAAAGATPACTGTCTGTAGATTGA	3157
Qy	641	ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg	660
Db	3158	CCAAACAGCTGGAATGGACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATAACAGA	3217
Qy	661	LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla	680
Db	3218	AAAGCCAAATCGGGTCAGAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTGCA	3277
Qy	681	AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu	700
Db	3278	GATAGGAAAGCTGTGATATCAACAGGAATGCTGAAATGTGTGTGTTCTTCAATGTTCTCCT	3337

QY 1181 GlyCysLeuIleSerPheIleIleIleSerTrpIleValArgLysAsnValAspThr 1200  
 Db 4778 GGTTCCTGATTTCTTTTCATTAAGATTTCTTGGAGATGTACGAAATAATGTGCACAC 4837  
 QY 1201 TyrAsnProTTPAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValIleu 1220  
 Db 4838 TATAATCATCGGATAGGCTTTTCAGTACTGTATATCGCCCTTACCTGCAGTGTGTACTG 4897  
 QY 1221 TrpIlePheLeuLeuGlnTyrTfGluIleIleIleSerTyrGlyArgSerIleArgLysAsp 1240  
 Db 4898 TGGATTTTCCTCTTCAACATACTATGAGAAATAATATGAGGAGCAGATCAATAAGAAAGAT 4957  
 QY 1241 ProPhePheArgAsnLeuSerThrIleSerLysAsnArgLysLeuProGluProProAsp 1260  
 Db 4958 CCTTTTTCAGAAACCTTTCAACGAGTCTAAATAATAGAAAGCTTTCAGAACCCACAGAC 5017  
 QY 1261 AsnGluAspGluAspGluAspValIleAlaGluArgLeuLysValLysGluLeuMetGly 1280  
 Db 5018 AATGAGGATGAAGATGAAGATGCTCAAGCTGAAAGACTAAAGGTCAAGAGCTGATGGGT 5077  
 QY 1281 CysGlnCysCysGluGluIleProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
 Db 5078 TGCCAGTGTGTGAGGAGAAACCATTCATTTATGGTCAGCAATTTGCATAAAGATATGAT 5137  
 QY 1301 AspIleLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer 1320  
 Db 5138 GACAGAAGATTTCTTTTCAGAAAGATGAAGAGTGGCAACTAAATACATCTCT 5197  
 QY 1321 PheCysValLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyIleSer 1340  
 Db 5198 TTCTGTGTGAAAGAAAGAGAGATCTTAGGACTATTGGGTCCAAATGTGTGGCAAGC 5257  
 QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
 Db 5258 ACATATTATATTTCTGTGTGTGTATTTGAACCACTTCAGCCAGGTATTTTAGGA 5317  
 QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
 Db 5318 GATTATTTCTTCAGACAGCAAGTGAAGATGATGATTCACCTGAAGTGTATGGGTACTGTCT 5377  
 QY 1381 GlnIleAsnProLeuTTPProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
 Db 5378 CAGATAAACCCCTTTGTGGCCAGATACATTCAGGAAACATTTTGAATTTTAGAGCT 5437  
 QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
 Db 5438 GTCAAGGAATGAGTGAAGTGAATGAAAGAGTCATAAGTCAATTAACACATGCACTT 5497  
 QY 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLys 1440  
 Db 5498 GATTTAAAGAACATCTTCAGAAAGACTGTAAAGAAACTACCTGCAGGAATCAACAGGAAG 5557  
 QY 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
 Db 5558 TTGTGTGTTCCTCTAAGTATGCTAGGGAATCCTCAGATTTCTTGTAGATGAACCATCT 5617  
 QY 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
 Db 5618 ACAGATGTGATCCCAAGCCCAACAGACATGTGGCAGCAATTCGAACTGCATTTAA 5677  
 QY 1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys 1500  
 Db 5678 AACAGAAAGCGGCTGCTATCTGACCACTCATATATGAGGAGGAGGCTGTCTGT 5737  
 QY 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
 Db 5738 GATCAGTAGCTATATGCTGTCTGGCAGTTAAGATGTATCGGAACAGTACACATCTA 5797  
 QY 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTyrIleGluAsn 1540  
 Db 5798 AAGATTAATTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGGAGTGTATAGAAAC 5857  
 QY 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560

Db 5858 CTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCAAATGCAAGCGTCA 5917  
 QY 1561 GluSerPheSerSerIleLeuAlaTyrIleProLysGluAspValGlnSerLeuSer 1580  
 Db 5918 GAAAGTTTTTCTTCTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTCA 5977  
 QY 1581 GlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluLysSerPhe 1600  
 Db 5978 CAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCCATTTGAAGAAATATAGCTTT 6037  
 QY 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
 Db 6038 TCTCAAGCAACATTCGAACAGGTTTGTGAGAACTCACTAAAGAAACAAGAGGAGAGAT 6097  
 QY 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyrTrpGluArgThrGlnGluAspArgVal 1640  
 Db 6098 AATAGTTGTGAACTTTTAAACAGCACTTTGTGGGAACGAAACAAGAGATAGAGTA 6157  
 QY 1641 ValPhe 1642  
 Db 6158 GTATTT 6163

RESULT 3

US-60-223-269-15  
 ; Sequence 15, Application US/60223269  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Yue, Henry  
 ; APPLICANT: Thornton, Michael  
 ; APPLICANT: Ramkumar, Jaya  
 ; APPLICANT: Tang, Y. Tom  
 ; APPLICANT: Azimzai, Yalda  
 ; APPLICANT: Baughn, Mariah R.  
 ; APPLICANT: Yang, Junming  
 ; APPLICANT: Yao, Monique G.  
 ; APPLICANT: Lal, Preeti  
 ; APPLICANT: Wallis, Marinder K.  
 ; APPLICANT: Gandhi, Ameena R.  
 ; APPLICANT: Hafalia, April  
 ; APPLICANT: Nguyen, Dannie B.  
 ; APPLICANT: Patterson, Chandra  
 ; APPLICANT: Elliott, Vicki S.  
 ; APPLICANT: Tribouley, Catherine M.  
 ; APPLICANT: Lu, Dyming Aina M.  
 ; APPLICANT: Xu, Yuming  
 ; APPLICANT: Reddy, Roopa  
 ; APPLICANT: Hernandez, Roberto  
 ; APPLICANT: Borowsky, Mark L.  
 ; APPLICANT: Lo, Terence P.  
 ; APPLICANT: Lu, Yan  
 ; APPLICANT: Policky, Jennifer L.  
 ; APPLICANT: Greene, Barrie D.  
 ; APPLICANT: Sanjanwala, Madhu S.  
 ; APPLICANT: Rauman, Brigitte E.  
 ; APPLICANT: Burford, Neil  
 ; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS  
 ; FILE REFERENCE: PI-0183 P  
 ; CURRENT APPLICATION NUMBER: US/60/223,269  
 ; CURRENT FILING DATE: 2000-08-03  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PERL Program  
 ; SEQ ID NO 15  
 ; LENGTH: 6369  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; OTHER INFORMATION: Incyte ID No: 2798241CB1  
 US-60-223-269-15

Alignment Scores: 0 Length: 6369  
 Pred. No.: 8422.00 Matches: 1641  
 Score:



Db 2558 TCAAGAGCAAGAAATATGAGGATTTATCAGAGGGCAATGTTAATGGAAATATTAGT 2617  
Qy 461 PheSerGluIleIleGluProValSerSerGluPheValGlyGlyGluAlaIleArgIle 480  
Db 2618 TTTAGTGAATATTAGGAGGAGTCTTCTTCTAGAAATTTGAGGAAAGAGGCATAGAAT 2677  
Qy 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
Db 2678 AGTGGTATTGAGAGACATACAGAAAGAGGGTGAATATGTGGAGCTTTGAGAAATTTG 2737  
Qy 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
Db 2738 TCATTTGACATATATGAGGGTACAGATTACTGCTTACTTGGCCACAGTGGACAGGAAG 2797  
Qy 521 SerThrLeuMetAsnIleLeuGlyGlyLeuGlyCysProSerAspGlyPheAlaSerIle 540  
Db 2798 AGTACATTGATGAATATCTTTGTGGACCTCTGCCACCTTCGTATGGGTTCATCTATA 2857  
Qy 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
Db 2858 TATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATGGCATT 2917  
Qy 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
Db 2918 TGTCCACAGTTAGATATACATTTGATGTTTGGACAGTAGAAGAAATTTATCAATTTG 2977  
Qy 581 AlaSerIleGlyGlyIleProLalaAsnIleIleGlnGluValGlnLysValLeuLeu 600  
Db 2978 GCTTCAATCAAGGGATACAGCAACAATATAATACAAAGAGTGCAGAGGTTTTACTA 3037  
Qy 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
Db 3038 GATTAGACATGACAGACTATCAAGATACCAAGCTAAATAATTAAGTGGTCAAAA 3097  
Qy 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuAspGlu 640  
Db 3098 AGAAGCTGTCTATGGAATTTGTGTTCTTTGGGAACCCAAAGATCTGCTGATGAA 3157  
Qy 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLysTyrArg 660  
Db 3158 CCAACAGCTGGATGGACCCCTGTTCTCGACATATTTGATGAATCTTTAAATACAGA 3217  
Qy 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
Db 3218 AAAGCAATCGGGTACAGAGTTCAGTACTCATTTTCATGGATCAAGTGCATCTTGCA 3277  
Qy 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 3278 GATAGGAAGCTGTGATATACAGGAATGCTGAAATGTTGGTCTTTCAATGTTCTTC 3337  
Qy 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 3338 AAAAGTAATGGGGATCGGTACCCGCTGAGCATGTACATAGACAAATATTGTGCCACA 3397  
Qy 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
Db 3398 GAATCTCTTTCTTCACTGGTTAAACACATATACCTGGAGTACTTTATTACACAGAAT 3457  
Qy 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 3458 GACCAACACTCTGTATAGTTCCTTCAAGGACATGACAAATTTTCAGGTTTGTT 3517  
Qy 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 3518 TCTGGCCCTACAGAGTCAATCAAAATTTGGGTGTCATTTCTATGGTGTTCCTATGACGACT 3577  
Qy 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 3578 TTGGAGACGCTATTTTAAGCTAGAGTTGAGCGAAGAAATTCACCAACAGATATAGT 3637  
Qy 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 3638 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTGATGAAATGGAA 3697

Qy 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 3698 CAGAGCTTACTTATTCTTTCTGAAACCAAGGCTTCTTAGTGAGCACCACATGAGCTTTGG 3757  
Qy 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db 3758 AAACAACAGATGTATCAATAGCAAGTTTCATTTCTTACCTTGAACGTAAGATAA 3817  
Qy 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 3818 TCAGTGAGATCAGTGTCTTCTGCTTTTAATTTTTCACAGTTTCAGATTTTATGTTT 3877  
Qy 881 LeuValHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 3878 TTGGTTTCATCAGTCTTTTAAATAATGTGTGGTTCCTCCATCAAACTTGTTCACAGCTTATAT 3937  
Qy 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
Db 3938 TTTCTAAACCTGGAGACAAACACATATAATACAAACAGAGTCTGCTTCTTCAAAATTTCT 3997  
Qy 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
Db 3998 GCTGACTCAGATATCAGTGTATCTTATAGCTTTTTCACAAAGCCAGACATAATGGTGAGC 4057  
Qy 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
Db 4058 ATGATTAATGACAGTGAATGATGATCGGTGGCTCCCATAGTGGCGCTTTAAATGTGGTG 4117  
Qy 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
Db 4118 CATTGAAAAGGACTATGTTTTTGAGCTGTTTTTCAACAGTACTATGTTTATTTCTTTA 4177  
Qy 981 ProIleLeuValAsnIleLeuSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
Db 4178 CCTATATTAGTAATATCATTAGTAATCTACTATCTTTTATCAITTTAAATGTGCTGAACC 4237  
Qy 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db 4238 ATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTTACTGATATAGTTTTTAAATTTGAG 4297  
Qy 1021 LeuTyrPheGlnAlaIleLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
Db 4298 CTGTATTTTCAAGCAGCTTTGCTTGGAACTATTGTTACTGCAATGCCACTTACTTTGCC 4357  
Qy 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 4358 ATGGAAATCGACAGAAATCATAGATCAAGCTTATACTCAACTTAAACTTTTCAGTCTTT 4417  
Qy 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleIle 1080  
Db 4418 TTGCCATCTGCATATTGGATTGGACAGCTGTTGTTGATATCCCTTATTTTATCAT 4477  
Qy 1081 LeuIleLeuMetLeuGlySerLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 4478 CTATTTTGTGCTAGGAAGCTTATGGCATTTTCAATATGATATATTTTATACTGTA 4537  
Qy 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
Db 4538 AAGTCTCTGCTGGTGTGTTTTTGCCTTATTTGTTTATGTTTCCATCAGTATTCTGTTCACT 4597  
Qy 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
Db 4598 TATATTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCAATTATC 4657  
Qy 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 4658 TATTCTGTGGCAGCGTTGGCTTGTATTGCAATCACTCACTGAAATAACTTTCTTTATGGGATAC 4717  
Qy 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleProIleTyrProLeuLeu 1180  
Db 4718 ACAATTGCAACTATTCTTCATTATGCTTTTGTATCATCATTCATTCATTCATTCATTCAT 4777

APPLICANT:	THORNTON, Michael; RAMKUMAR, Jayalaxmi;
APPLICANT:	TANG, Y. Ton; AZIMZAI, Yalda;
APPLICANT:	BAUGHN, Mariah R.; YANG, Junming;
APPLICANT:	YAO, Monique G.; LAL, Preeti G.;
APPLICANT:	CHAWLA, Narinder K.; GANDHI, Ameena R.;
APPLICANT:	HAFALIA, April J.A.; NGUYEN, Dannel B.;
APPLICANT:	ARVIZU, Chandra S.; ELLIOTT, Vicki S.;
APPLICANT:	TRIBOULEY, Catherine M.; LU, Dyung Aina M.;
APPLICANT:	XU, Yuming; REDDY, Roopa;
APPLICANT:	HERNANDEZ, Roberto; BOROWSKY, Mark L.;
APPLICANT:	LO, Terence P.; LU, Yan;
APPLICANT:	POLICKY, Jennifer L.; GREENE, Barrie D.;
APPLICANT:	SANJANWALA, Madhusudan M.; RAUMANN, Brigitte E.;
APPLICANT:	BURFORD, Neil; ISON, Craig H.;
APPLICANT:	LEE, Ernestine A.; DING, Li.;
APPLICANT:	DAS, Debopriya; KALLICK, Deborah A.;
APPLICANT:	KHAN, Farrah A.; SEILHAMER, Jeffrey J.;
TITLE OF INVENTION:	TRANSPORTERS AND ION CHANNELS
FILE REFERENCE:	PI-0183 USN
CURRENT APPLICATION NUMBER:	US/10/343,903
CURRENT FILING DATE:	2003-02-03
PRIOR APPLICATION NUMBER:	PCT/US01/24217
PRIOR FILING DATE:	2001-08-01
PRIOR APPLICATION NUMBER:	60/231,434
PRIOR FILING DATE:	2000-09-08
PRIOR APPLICATION NUMBER:	60/230,067
PRIOR FILING DATE:	2000-08-31
PRIOR APPLICATION NUMBER:	60/228,140
PRIOR FILING DATE:	2000-08-25
PRIOR APPLICATION NUMBER:	60/226,410
PRIOR FILING DATE:	2000-08-18
PRIOR APPLICATION NUMBER:	60/224,456
PRIOR FILING DATE:	2000-08-10
PRIOR APPLICATION NUMBER:	60/223,269
PRIOR FILING DATE:	2000-08-03
NUMBER OF SEQ ID NOS:	60
SOFTWARE:	PERL Program
SEQ ID NO 33	
LENGTH:	6369
TYPE:	DNA
ORGANISM:	Homo sapiens
FEATURE:	
NAME/KEY:	misc feature
OTHER INFORMATION:	Incbyte ID No: 2798241CB1
US-10-343-903-33	
Alignment Scores:	
Pred No.:	0
Length:	6369
Score:	8422.00
Percent Similarity:	100.00%
Conservative:	1
Best Local Similarity:	99.94%
Mismatches:	0
Query Match:	99.95%
Indels:	0
Gaps:	0
DB:	
US-10-090-458-5 (1-1642) x US-10-343-903-33 (1-6369)	
Qy	1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20
Db	1238 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGAGACACACACAGACACTTCTACTGAAG 1297
Qy	21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGlnIleLeuPheProLeu 40
Db	1298 AATTACTTAATTAATGACAGACCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 1357
Qy	41 PhePheLeuPheTrpLeuLeuLeuLeuLeuSerMetMetHisProAsnLysLysTrpGluGlu 60
Db	1358 TTTTITTTTATTTTGGTTAATATTAATAGCATGATGATCCCAATAAGAAATATGAAGAA 1417
Qy	61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80
Db	1418 GTGCCCAATATAGAACTCAATCTCTATGACAGAGTGTACTCTTTCTTAATCTTCTTGG 1477
Qy	81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100

QY	961	HisSerGluLysAspTyrValPheAlaValPheAsnSerThrMetValTyrSerLeu	980
DB	2994	CATTTCAGAAAGGACTATGTTTTCAGCTGTTTTCACAGPACTATGGTTATTTCTTTA	3053
QY	981	ProLeuValAsnIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr	1000
DB	3054	CCTATATTAGTGAATATCATTAAGTAACTATTTTAAATGTGACTGAAACC	3113
QY	1001	IleGlnIleTyrSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu	1020
DB	3114	ATCCAGATCTGAGTACCCCATCTTTTCAAGAAATTTACTGATATAGTTTAAATAGAG	3173
QY	1021	LeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla	1040
DB	3174	CTGTATTTTCAGCAGCTTGTGGAATCAATGTTACTGCAATGCCACCTTACTTGGC	3233
QY	1041	MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerClyLeu	1060
DB	3234	ATGGAAATTCGAGAAATCATTAAGATCAAGAGCTTACTCAACTTAAACTTTCAGGTCTT	3293
QY	1061	LeuProSerAlaTyrTyrIleGlyGlnAlaValAspIleProLeuPhePheIleIle	1080
DB	3294	TTCCTATCTGCATTTGGATTGGACAGCTGTTGTTGATATCCCTTATTTTATCAT	3353
QY	1081	LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal	1100
DB	3354	CTTATTTTCATCTAGCAAGCTTATGGCATTTTCAATATGATGATTTATTTTATCTGA	3413
QY	1101	LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr	1120
DB	3414	AGTTCTCTGCTGCTGTTTTCCTTATGCTTATGTTATGTTCCATCAGTTATCTGTTCACT	3473
QY	1121	TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTyrPheIle	1140
DB	3474	TATATTGCTTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGGTCATTTATC	3533
QY	1141	TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr	1160
DB	3534	TATCTGTGGCAGCTGCTGCTGTTATGCAATCACTGAATTAACCTTCTTTATGGGATAC	3593
QY	1161	ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu	1180
DB	3594	ACAAATTCGAATCTTCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT	3653
QY	1181	GlyCysLeuIleSerPheIleLysIleSerTyrLysAsnValArgLysValAspThr	1200
DB	3654	GGTTGCTGATTTCTTTTCAAGATTTCTTGAAGAATGTACGAAATTTTGGACACC	3713
QY	1201	TyrAsnProTyrAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu	1220
DB	3714	TATAATCCATGGATAGGCTTTTCAGTATGCTTATGCTTATGCTTATGCTTATGCTTATGCT	3773
QY	1221	TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyLysArgSerIleArgLysAsp	1240
DB	3774	TGGAATTTCTCTCTTCAATACATATGAGAAATATATGGAGGAGATCAATAGAAAGAT	3833
QY	1241	ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAsp	1260
DB	3834	CCCTTTTTCAGAAACCTTTTCAGAGTCTTAAATATAGAACTTCCAGAACCCACAGAC	3893
QY	1261	AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly	1280
DB	3894	AATGAGGATGAAGATGAAGATGTCAAAGCTGAAAGCTTAAAGCTTAAAGCTTAAAGCT	3953
QY	1281	CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp	1300
DB	3954	TGCAGTGTGTGAGAGAAACCATTCATTTATGCTTATGCTTATGCTTATGCTTATGCT	4013
QY	1301	AspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer	1320
DB	4014	GACAAGAAAGATTTCTTCTTTTCAGAAAGTAAAGAAAGTGGCACTAAATATCATCTCT	4073

QY	1321	PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer	1340
DB	4074	TTCTGTGTGAAAACAGAGAGATCTTTAGGACTATTGGGTCCAAATGGTCTGGCAAAAC	4133
QY	1341	ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerClyGlnValPheLeuGly	1360
DB	4134	ACAATATTAAATATTCTGTTGTTGATATTGAACCAACTTCAGCCAGGATTTTTTAGA	4193
QY	1361	AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro	1380
DB	4194	GATATTCTTCAGAGCAAGTGAAGATGATTCATCTGAAGTGTATGGTTACTCTCT	4253
QY	1381	GlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla	1400
DB	4254	CAGATAAACCTTTTGTGGCAGATACATATTGCAGGAACATTTTGAAATTTATGGAGCT	4313
QY	1401	ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu	1420
DB	4314	GTCAAGGAATGATGTCAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG	4373
QY	1421	AspLeuLysGluHisLeuGlnLysThrValLysLeuProAlaGlyLysArgLys	1440
DB	4374	GATTTAAAGAACATCTTCAGAGACTGTAAGAAACTACCTGCAGGAATCAACAGAAAG	4433
QY	1441	LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer	1460
DB	4434	TTGTGTTTCTCTTAAGTATGCTAGGAATCTTCAGATTTACTTTGCTAGTGAACCATCT	4493
QY	1461	ThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLys	1480
DB	4494	ACAGGTATGATCCCAAGCAACAGCACATGTGGCAGCAATTCGAACTGCATTTAA	4553
QY	1481	AsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCys	1500
DB	4554	AAAGAAAGCGGCTGCTATCTTCAGCACCTCATATATGAGGAGGAGGAGGCTGCTCTGT	4613
QY	1501	AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu	1520
DB	4614	GATCGAGTAGCTATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	4673
QY	1521	LysSerLysPheGlyLysGlyTyrPheLeuGluLysLeuLysLeuLysLeuLysLeuLys	1540
DB	4674	AAAGATTAATTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGGACTGGATAGAAAC	4733
QY	1541	LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln	1560
DB	4734	CTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATTCAGAGCGCTCAG	4793
QY	1561	GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer	1580
DB	4794	GAAGATTTTCTCTTATTTTGGCTTATAAATTCCTTAAGGAAGATTTTTCAGTCCCTTTCA	4853
QY	1581	GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe	1600
DB	4854	CAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGGCATTCAGAAATATAGCTTT	4913
QY	1601	SerGlnAlaThrLeuGlnValPheValGluLeuThrLysGluGlnGluGluAsp	1620
DB	4914	TCTCAGCAACATTTGGAACAGGTTTGTAGAACTCCTCAAAAGCAAGAGGAGGAGAT	4973
QY	1621	AsnSerCysGlyThrLeuAsnSerThrLeuTyrTyrGluArgThrGlnGluAspArgVal	1640
DB	4974	AATAGTTGTGAACTTTTAAACAGCACATCTTTGTGGAGCAACACAGAGAGATAGAT	5033
QY	1641	ValPhe 1642	
DB	5034	GTAATTT 5039	

RESULT 2  
US-10-343-903-33  
; Sequence 33, Application US/10343903  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.; YUE, Henry;

Db	774	TATATACCTAGTTATAGCATTTTCCCTTTTGGATACCTTTTGGCAATTCATATC	833
Qy	241	ValAlaGluLysGluLysGluLysGluPheLeuLysIleMetGlyLeuHisAspThr	260
Db	834	GTAGCAGAAAAGAAAAATAAAGAAATTTTAAAGATAAATGGGACTTCATGACT	893
Qy	261	AlaPheTyrLeuSerTyrValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu	280
Db	894	GCCCTTTGGCTTCTCTGGTTCTTCTATATACAGTTTAATTTTCTTATGTCCTCTCT	953
Qy	281	MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe	300
Db	954	ATGGCAGTCATATGGGACAGCTTCTTTGTTATTTCTCTCAAGTAGCAGCATTTGGATATT	1013
Qy	301	LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu	320
Db	1014	CTGCTTTTTCCTTTATGGATATCATCTGTAATTTTGTCTTAAATGCTGACACCTCTT	1073
Qy	321	PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe	340
Db	1074	TTTAAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTGTACTGTGGCTTTTGGATTT	1133
Qy	341	IleGlyLeuMetIleLeuLeuLeuSerPheProLysSerLeuValTyrLeuPheSer	360
Db	1134	ATTGGCTTATGATAATCTCTCATGAAGATTTTCCAAATCGTTAGTGTGGCTTTTCAGT	1193
Qy	361	ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe	380
Db	1194	CCTTCTGTCACTGTACTTTTGTGATTTGGTATTTGCATGAGGTGATGCAATTTAGAAGATTT	1253
Qy	381	AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle	400
Db	1254	AATGAAGTGCTTCATTTTCAAAATTTGACTGCGAGCCCATATCTCTTAATTTATACAAAT	1313
Qy	401	IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal	420
Db	1314	ATCATGCTCACATTAATAGTATATTTCTATGTCCTTTGGCTGTCTATCTTGATCAAGTC	1373
Qy	421	IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTyr	440
Db	1374	ATTCAGGGGAATTTGGCTTACGAGATCATCTTTATATTTTCTGAAGCCCTTCATATGG	1433
Qy	441	SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer	460
Db	1434	TCAAAGACAAAAGAAATATAGGAGTTATCAGAGGCAATGTTATGGAAATATAGT	1493
Qy	461	PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle	480
Db	1494	TTTAGTGAATATATGAGCCAGTTTCTTCAAGATTTGTAGGAAAAGAGCCATAAGAATT	1553
Qy	481	SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu	500
Db	1554	AGTGGTATTCAGAAGACATACAGAAAGAGGGTGAAAATGTGGAGGCTTTGAGAAATTTG	1613
Qy	501	SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys	520
Db	1614	TCATTTGATATATAGGGTTCAGATTACTGCCTTACTTGGCCACAGTGGAAACAGGAAG	1673
Qy	521	SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle	540
Db	1674	AGTACATATGATATATCTTTTGGACTCTGCCACCTCTCTGATGGGTTTGATCTATA	1733
Qy	541	TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle	560
Db	1734	TATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAGAGAAAATGATTTGGCAT	1793
Qy	561	CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu	580
Db	1794	TGTCACAGTTAGATATACATTTGATGTTTGGAGTAGAAGAAAATTTATCAATTTTG	1853
Qy	581	AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu	600

Db	1854	GCTTCAATCAAAAGGATACAGCCCAACATATAATACAAAGAGTGACAGAGGTTTTACTA	1913
Qy	601	AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys	620
Db	1914	GAATTAGACATCGACATATCAAGATTAACCAAGCTAAAAAATTAAGTGGTGTCAAAA	1973
Qy	621	ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu	640
Db	1974	AGAAAGCTGTCAATAGGAATGCTGTCTTGGAAACCAAAAGATATCTGCTGTAGTAA	2033
Qy	641	ProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLeuLysTyrArg	660
Db	2034	CAAAAGCTGGGAATGGACCCCTGTTCTGCACATATGATGGAATCTTTTAAAAATACAG	2093
Qy	661	LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla	680
Db	2094	AAAGCCAAATCGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTG	2153
Qy	681	AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu	700
Db	2154	GATAGAAAGCTGTGATATCACAGGAATGCTGAAATGTGTGGTTCTTCAATGTCTCTC	2213
Qy	701	LysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr	720
Db	2214	AAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACA	2273
Qy	721	GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrIleLeuGlnIleAsn	740
Db	2274	GAATCTCTTCTCTCACTGCTTAAACAACATATACCTGGAGCTACTTTATTACACAGAA	2333
Qy	741	AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe	760
Db	2334	GAACCAACACTTGTGTATAGCTTGGCTTTCAGGACATGGACAAATTTTCAGGTGTGTT	2393
Qy	761	SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr	780
Db	2394	TCTGGCCCTAGACAGTCATTTCAAAATTTGGGTGTCATTTCTTATGTTGTTTCCATG	2453
Qy	781	LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer	800
Db	2454	TTGGAAGAGCTATTTTAAAGCTAGAGTTGAAGCAGAAATGACCAAGCATTTAGT	2513
Qy	801	ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu	820
Db	2514	GTATTTACTCAGCAGCACTGGAGGAAGAAATGGATTCAAAATCTTTTGATGAAATGAA	2573
Qy	821	GlnSerLeuLeuIleLeuSerGlnThrLysAlaSerLeuValSerThrMetSerLeuThr	840
Db	2574	CAGAGCTTACTTATCTTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTTG	2633
Qy	841	LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys	860
Db	2634	AAACAACAGATCTATACAAATAGCAAAAGTTTCATTTCTTACCTTGAAACGTGAAAGTAA	2693
Qy	861	SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe	880
Db	2694	TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAGTTTCAGATTTTATGTT	2753
Qy	881	LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr	900
Db	2754	TTGGTTTCACTCTTTTAAAAATGCTGTGGTTCCCATCAAACTTTGTTCCAGCTTATAT	2813
Qy	901	PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer	920
Db	2814	TTTCTAAAAACCTGGAGACAAACCAATATAAACAACCAAGTCTCTTCTTCAAAATTTCT	2873
Qy	921	AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr	940
Db	2874	GCTGACTCAGATATCAGTGATCTTTATTAGCTTTTTCACAGCCAGAACATATGTTGAG	2933
Qy	941	MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet	960
Db	2934	ATGATTAATGACAGTGACTATGTAACCGTGGCTCCCATAGTGGCTTTAAATGTGATG	2993

102: /cgn2\_6/prodata/2/pna/us6045\_COMB.seq:  
103: /cgn2\_6/prodata/2/pna/us6046\_COMB.seq:  
104: /cgn2\_6/prodata/2/pna/us6047\_COMB.seq:  
105: /cgn2\_6/prodata/2/pna/us6048\_COMB.seq:  
106: /cgn2\_6/prodata/2/pna/us6049\_COMB.seq:  
107: /cgn2\_6/prodata/2/pna/us6050\_COMB.seq:  
108: /cgn2\_6/prodata/2/pna/us6051\_COMB.seq:  
109: /cgn2\_6/prodata/2/pna/us6052\_COMB.seq:  
110: /cgn2\_6/prodata/2/pna/us6053\_COMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	ID	Description
1	8426	100.0	5475	44	US-10-090-458-4
2	8422	100.0	6369	49	US-10-343-903-33
3	8422	100.0	6369	78	US-60-223-269-15
4	8422	100.0	6525	43	US-10-005-338B-1
5	8395	99.6	4929	41	US-09-971-121-1
6	8395	99.6	5262	41	US-09-971-121-5
7	8309	98.6	4917	44	US-10-090-458-3
8	8309	98.6	5463	44	US-10-090-458-1
9	8253.5	98.0	5004	1	PCT-US03-33087-82
10	8112	95.3	4785	41	US-09-971-121-3
11	8028.5	95.3	6170	1	PCT-US03-14026-75
12	7377	87.6	5595	102	US-60-453-050-2441
13	7377	87.6	5595	102	US-60-453-135-2441
14	7377	87.6	5595	103	US-60-466-412-2441
15	4930.5	58.5	3347	45	US-10-108-260A-160
16	4847.5	57.5	3268	44	US-10-094-749-984
17	4382.5	52.0	3448	45	US-10-144-771-16659
18	4382.5	52.0	3448	93	US-60-360-207-16659
19	4311	51.2	3368	74	US-60-185-213-1454
20	4272	50.7	2573	78	US-60-226-176-242
21	4272	50.7	2573	79	US-60-233-468-242
22	4272	50.7	2573	88	US-60-313-371-241
23	4265	50.6	2926	74	US-60-185-213-1628
24	4265	50.6	2926	84	US-60-278-232-4057
25	3672	43.6	2887	78	US-60-226-176-241
26	3672	43.6	2887	79	US-60-233-468-241
27	3672	43.6	2887	88	US-60-313-371-241
28	3656	43.4	2988	78	US-60-226-176-237
29	3656	43.4	2988	79	US-60-233-468-237
30	3656	43.4	2988	88	US-60-313-371-237
31	3654	43.4	2109	78	US-60-226-176-240
32	3654	43.4	2109	79	US-60-233-468-240
33	3654	43.4	2109	88	US-60-313-371-240
34	3283.5	39.0	2502	20	US-09-496-914A-3176
35	3283.5	39.0	2502	24	US-09-560-875A-3176
36	3266	38.8	4302	1	PCT-US03-30720-38
37	3259	38.7	5797	1	PCT-US02-33496-1
38	3259	38.7	5797	1	PCT-US03-28227-2563
39	3234	38.4	5445	1	PCT-US03-33087-108
40	3234	38.4	6181	43	US-10-005-338B-4
41	3229	38.3	5722	46	US-10-171-423-3
42	3200.5	38.0	4875	44	US-10-090-454-3
43	3200.5	38.0	5018	44	US-10-090-454-1
44	3197.5	37.9	5981	43	US-10-005-338B-3
45	3195	37.9	5311	1	PCT-US02-41825-88

ALIGNMENTS

RESULT 1  
US-10-090-458-4  
; Sequence 4, Application US/10090458  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Hongyun

; APPLICANT: Killinski, Ligia  
; APPLICANT: Le Bihan, Stephane  
; TITLE OF INVENTION: NOVEL ABCAS TRANSPORTER AND USES THEREOF  
; FILE REFERENCE: 100103.403  
; CURRENT APPLICATION NUMBER: US/10/090,458  
; CURRENT FILING DATE: 2002-03-01  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 5475  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 25  
; OTHER INFORMATION: n = A,T,C or G  
US-10-090-458-4

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
0	8426.00	5475	1642
100.00%	Conservative:	0	
100.00%	Mismatches:	0	
100.00%	Indels:	0	
44	Gaps:		
US-10-090-458-5 (1-1642) x US-10-090-458-4 (1-5475)			
Qy	1	MetSerThrAlaIleArgGluValTpaArgGlnThrArgThrLeuLeuLys	20
Db	114	ATGTCACCTGCATTAAGGAGGTAGGAGTTGGACACAGACCACTTCTACTGAAG	173
Qy	21	AsnTyrLeuIleLysCysArgThrLysSerValGlnGluLeuLeuPheProLeu	40
Db	174	AATTACTTAATTAATCGAGAACCAAAAGAGTAGTGTTCAGGAAATCTTTTCCACTA	233
Qy	41	PhePheLeuPheThrPheLeuLeuLeuSerMetMetHisProAsnLysLysTyrGluGlu	60
Db	234	TTTTTTTTTTTGGTATATTAATACGATGATGCAATCCAAATAGAAATATGAGAA	293
Qy	61	ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly	80
Db	294	GTGCTTAATATAGAACTCAATCTATGGACAAAGTTTACTCTTCTTAATCTAATCTTGA	353
Qy	81	TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu	100
Db	354	TATACTCCAGTGAATTAATTAACAGCAGCATCGAGAAAGTGCTACTGATCATCTA	413
Qy	101	ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu	120
Db	414	CCTGATGTCATTAATTAATGAGATATACAAATGAAAGAAATGTTAACAATCCAGTCTC	473
Qy	121	SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg	140
Db	474	TCTAAGCCGAGCAACTTTGAGTGTGGTTTTCAAGACTCCATGCTCTATGAATCTGT	533
Qy	141	PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer	160
Db	534	TTTTTTCCTGATATGATTCAGTATCTTCTATTATATGATTCAGAGCTGGCTGTTC	593
Qy	161	LysSerCysGluAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnIleSerIle	180
Db	594	AAATCATGTGAGGCTGCTCAGTACTGGTCTCTCAGGTTTCACAGTTTACAAGCATCCATA	653
Qy	181	AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGluSerThr	200
Db	654	GATGTCGCCATTATACAGTTGAAGCAATGTTTCTCTTTGGAGAGGCTGAGTCACT	713
Qy	201	LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal	220
Db	714	AAAGCTGTATTATGGAGAACTGCTGTGTAGAAATAGATACCTTTCCCGAGGAGTA	773
Qy	221	IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle	240

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 12, 2004, 21:29:55 ; Search time 7298 Seconds  
(without alignments)

7916.636 Million cell updates/sec

Title: US-10-090-458-5

Perfect score: 8426

Sequence: 1 MSTAIREVGVNRQRTLLK.....CGTLNSTLWERTQEDRWVF 1642

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 37577330 seqs, 17593059518 residues

Total number of hits satisfying chosen parameters: 75154660

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US10090458/runat\_08042004\_121553\_25595/app\_query.fasta\_1.1799  
-DB=Pending Patents NA.Main -QWTF=fastap -SUFFIX=rpnm -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFW=ptp -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10090458@cgn\_1.1\_4733 -runat\_08042004\_121553\_25595 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending Patents NA.Main:  
1: /cgn2\_6/ptodata/2/pna/US06\_PCTUS\_COMB.seq.\*  
2: /cgn2\_6/ptodata/2/pna/US06\_COMB.seq.\*  
3: /cgn2\_6/ptodata/2/pna/US07\_COMB.seq.\*  
4: /cgn2\_6/ptodata/2/pna/US080\_COMB.seq.\*  
5: /cgn2\_6/ptodata/2/pna/US081\_COMB.seq.\*  
6: /cgn2\_6/ptodata/2/pna/US082\_COMB.seq.\*  
7: /cgn2\_6/ptodata/2/pna/US083\_COMB.seq.\*  
8: /cgn2\_6/ptodata/2/pna/US084\_COMB.seq.\*  
9: /cgn2\_6/ptodata/2/pna/US085\_COMB.seq.\*  
10: /cgn2\_6/ptodata/2/pna/US086\_COMB.seq.\*  
11: /cgn2\_6/ptodata/2/pna/US087\_COMB.seq.\*  
12: /cgn2\_6/ptodata/2/pna/US088\_COMB.seq.\*  
13: /cgn2\_6/ptodata/2/pna/US089\_COMB.seq.\*  
14: /cgn2\_6/ptodata/2/pna/US090\_COMB.seq.\*  
15: /cgn2\_6/ptodata/2/pna/US091\_COMB.seq.\*  
16: /cgn2\_6/ptodata/2/pna/US092A\_COMB.seq.\*  
17: /cgn2\_6/ptodata/2/pna/US092B\_COMB.seq.\*  
18: /cgn2\_6/ptodata/2/pna/US093A\_COMB.seq.\*  
19: /cgn2\_6/ptodata/2/pna/US093B\_COMB.seq.\*  
20: /cgn2\_6/ptodata/2/pna/US094\_COMB.seq.\*  
21: /cgn2\_6/ptodata/2/pna/US095A\_COMB.seq.\*  
22: /cgn2\_6/ptodata/2/pna/US095B\_COMB.seq.\*  
23: /cgn2\_6/ptodata/2/pna/US095C\_COMB.seq.\*  
24: /cgn2\_6/ptodata/2/pna/US095D\_COMB.seq.\*  
25: /cgn2\_6/ptodata/2/pna/US096A\_COMB.seq.\*  
26: /cgn2\_6/ptodata/2/pna/US096B\_COMB.seq.\*  
27: /cgn2\_6/ptodata/2/pna/US096C\_COMB.seq.\*  
28: /cgn2\_6/ptodata/2/pna/US096D\_COMB.seq.\*

29: /cgn2\_6/ptodata/2/pna/US096E\_COMB.seq.\*  
30: /cgn2\_6/ptodata/2/pna/US097A\_COMB.seq.\*  
31: /cgn2\_6/ptodata/2/pna/US097B\_COMB.seq.\*  
32: /cgn2\_6/ptodata/2/pna/US097C\_COMB.seq.\*  
33: /cgn2\_6/ptodata/2/pna/US098A\_COMB.seq.\*  
34: /cgn2\_6/ptodata/2/pna/US098B\_COMB.seq.\*  
35: /cgn2\_6/ptodata/2/pna/US098C\_COMB.seq.\*  
36: /cgn2\_6/ptodata/2/pna/US098D\_COMB.seq.\*  
37: /cgn2\_6/ptodata/2/pna/US099A\_COMB.seq.\*  
38: /cgn2\_6/ptodata/2/pna/US099B\_COMB.seq.\*  
39: /cgn2\_6/ptodata/2/pna/US099C\_COMB.seq.\*  
40: /cgn2\_6/ptodata/2/pna/US099D\_COMB.seq.\*  
41: /cgn2\_6/ptodata/2/pna/US099E\_COMB.seq.\*  
42: /cgn2\_6/ptodata/2/pna/US099F\_COMB.seq.\*  
43: /cgn2\_6/ptodata/2/pna/US100A\_COMB.seq.\*  
44: /cgn2\_6/ptodata/2/pna/US100B\_COMB.seq.\*  
45: /cgn2\_6/ptodata/2/pna/US101A\_COMB.seq.\*  
46: /cgn2\_6/ptodata/2/pna/US101B\_COMB.seq.\*  
47: /cgn2\_6/ptodata/2/pna/US102A\_COMB.seq.\*  
48: /cgn2\_6/ptodata/2/pna/US102B\_COMB.seq.\*  
49: /cgn2\_6/ptodata/2/pna/US103A\_COMB.seq.\*  
50: /cgn2\_6/ptodata/2/pna/US103B\_COMB.seq.\*  
51: /cgn2\_6/ptodata/2/pna/US104A\_COMB.seq.\*  
52: /cgn2\_6/ptodata/2/pna/US104B\_COMB.seq.\*  
53: /cgn2\_6/ptodata/2/pna/US106\_COMB.seq.\*  
54: /cgn2\_6/ptodata/2/pna/US107A\_COMB.seq.\*  
55: /cgn2\_6/ptodata/2/pna/US107B\_COMB.seq.\*  
56: /cgn2\_6/ptodata/2/pna/US6000\_COMB.seq.\*  
57: /cgn2\_6/ptodata/2/pna/US6001\_COMB.seq.\*  
58: /cgn2\_6/ptodata/2/pna/US6002\_COMB.seq.\*  
59: /cgn2\_6/ptodata/2/pna/US6003\_COMB.seq.\*  
60: /cgn2\_6/ptodata/2/pna/US6004\_COMB.seq.\*  
61: /cgn2\_6/ptodata/2/pna/US6005\_COMB.seq.\*  
62: /cgn2\_6/ptodata/2/pna/US6006\_COMB.seq.\*  
63: /cgn2\_6/ptodata/2/pna/US6007\_COMB.seq.\*  
64: /cgn2\_6/ptodata/2/pna/US6008\_COMB.seq.\*  
65: /cgn2\_6/ptodata/2/pna/US6009\_COMB.seq.\*  
66: /cgn2\_6/ptodata/2/pna/US6010\_COMB.seq.\*  
67: /cgn2\_6/ptodata/2/pna/US6011\_COMB.seq.\*  
68: /cgn2\_6/ptodata/2/pna/US6012\_COMB.seq.\*  
69: /cgn2\_6/ptodata/2/pna/US6013\_COMB.seq.\*  
70: /cgn2\_6/ptodata/2/pna/US6014\_COMB.seq.\*  
71: /cgn2\_6/ptodata/2/pna/US6015\_COMB.seq.\*  
72: /cgn2\_6/ptodata/2/pna/US6016\_COMB.seq.\*  
73: /cgn2\_6/ptodata/2/pna/US6017\_COMB.seq.\*  
74: /cgn2\_6/ptodata/2/pna/US6018\_COMB.seq.\*  
75: /cgn2\_6/ptodata/2/pna/US6019\_COMB.seq.\*  
76: /cgn2\_6/ptodata/2/pna/US6020\_COMB.seq.\*  
77: /cgn2\_6/ptodata/2/pna/US6021\_COMB.seq.\*  
78: /cgn2\_6/ptodata/2/pna/US6022\_COMB.seq.\*  
79: /cgn2\_6/ptodata/2/pna/US6023A\_COMB.seq.\*  
80: /cgn2\_6/ptodata/2/pna/US6023B\_COMB.seq.\*  
81: /cgn2\_6/ptodata/2/pna/US6024\_COMB.seq.\*  
82: /cgn2\_6/ptodata/2/pna/US6025\_COMB.seq.\*  
83: /cgn2\_6/ptodata/2/pna/US6026\_COMB.seq.\*  
84: /cgn2\_6/ptodata/2/pna/US6027\_COMB.seq.\*  
85: /cgn2\_6/ptodata/2/pna/US6028\_COMB.seq.\*  
86: /cgn2\_6/ptodata/2/pna/US6029\_COMB.seq.\*  
87: /cgn2\_6/ptodata/2/pna/US6030\_COMB.seq.\*  
88: /cgn2\_6/ptodata/2/pna/US6031\_COMB.seq.\*  
89: /cgn2\_6/ptodata/2/pna/US6032\_COMB.seq.\*  
90: /cgn2\_6/ptodata/2/pna/US6033\_COMB.seq.\*  
91: /cgn2\_6/ptodata/2/pna/US6034\_COMB.seq.\*  
92: /cgn2\_6/ptodata/2/pna/US6035\_COMB.seq.\*  
93: /cgn2\_6/ptodata/2/pna/US6036\_COMB.seq.\*  
94: /cgn2\_6/ptodata/2/pna/US6037\_COMB.seq.\*  
95: /cgn2\_6/ptodata/2/pna/US6038\_COMB.seq.\*  
96: /cgn2\_6/ptodata/2/pna/US6039\_COMB.seq.\*  
97: /cgn2\_6/ptodata/2/pna/US6040\_COMB.seq.\*  
98: /cgn2\_6/ptodata/2/pna/US6041\_COMB.seq.\*  
99: /cgn2\_6/ptodata/2/pna/US6042\_COMB.seq.\*  
100: /cgn2\_6/ptodata/2/pna/US6043\_COMB.seq.\*  
101: /cgn2\_6/ptodata/2/pna/US6044\_COMB.seq.\*

```
QY 1138 SerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleIleThrGluIleThrPhePhe 1157
DDB 3391 TCATTTTACTCTCTTTTGGCTCCACCATCATGTTTCCATCATCTTAATCAATCATTTT 3450
QY 1158 ---MetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIle 1176
DDB 3451 GACCTAAGTATATTGATTACCACCATG-----GTATTGGTTCTCTCA 3492
QY 1177 TyrProLeuLeuGlyCysLeuIleSerPheIleIleIleIleSerTyrPheAsnValArgLys 1196
DDB 3493 TATACCTTGGCTTGA---TTTAAACCTTTTGGAGTG-----AGA 3531
QY 1197 AsnValAspThrTyrAsnProTyrPheArgLeuSer 1208
DDB 3532 GACCAGGAGCAGCTACAGAGAATTTCCAGAGCAAAATTTGAATTTGAGTGCCACTGATTTT 3591
QY 1209 ValAlaValIleSerProTyrLeuGlnCysValLeuTyrIlePheLeuLeuGlnTyrTyr 1228
DDB 3592 CTAGTCTGCTTCATACCTCTTTTCAGACTTTGCTATTCGTTTGTCTTAAGATGCGATG 3651
QY 1229 GluLysIleTyrGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThr 1248
DDB 3652 GAACTAAATGTGGAAGAAAGAAATGCAAAAGATCTGTTTTCAGA---ATTTCCTCC 3708
QY 1249 LysSerLysAsnArgLysLeuProGluProProAspAsnGluAspGluAspVal 1268
DDB 3709 CAAGTAGAGATGCTAAG---CCAAATCCAGAGACCCATAGATGAAGATGAATATT 3765
QY 1269 LysAlaGluArgLeuLysValLysValLysLeuMetGlyCysGlnCysCysGluGluLysPro 1288
DDB 3766 CAACAGAGAAAGATAAGAACAGCCACTGCTCTGACCACTTCAATCTTAGATGAGAAACCT 3825
QY 1289 SerIleMetValSerAsnLeuHisLysGluTyrAspLysLysAspPheLeuLeuSer 1308
DDB 3826 GTTATAATGGCAGCTGTCTACAAAGAAATATCGAGCCAGAGAAAGTTGCTTTTCA 3885
QY 1309 ArgLysValLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIle 1328
DDB 3886 AAGAGAGAGAGAGAAATAGCAGCAGAGAAATATCTTCTGTTCTCAAGAGGTGAATTT 3945
QY 1329 LeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThrIleIleAsnIleLeuValGly 1348
DDB 3946 TTGGGATTCAGGACCCAGTGGTGTGGAAGAAAGTTCACTATTAGAATGATATCTGGG 4005
QY 1349 AspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGlu 1368
DDB 4006 ATCAAAAGCCCACTGCTGGAGAGGTGAATGAAAGCTGCAGTTCA----- 4053
QY 1369 AspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTyrProAsp 1388
DDB 4054 -----GTTTGGGCCACTGGGGTACTGCGCTCAAGAGAACGTGCTGTGGCCCATG 4104
QY 1389 ThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAsp 1408
DDB 4105 CTGAGTTGAGGGAGAACCTGGAGGTGATCTGCGCTCAAGGGGCTCAGGAAGCGGAC 4164
QY 1409 MetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLys 1428
DDB 4165 GCGAGGCTCGCCATCGCAAGATTAGTGAGTGCTTTCAAACTGCATGAGCAGCTGAATGTT 4224
QY 1429 ThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeu 1448
DDB 4225 CCTGTCAGAAATTAACAGCAGGAATCACGAGAAGTTGTGTTTGTGCTGAGCCTCTG 4284
QY 1449 GlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLys 1468
DDB 4285 GGAAGACTCACCTGCTCTCTGATGACCACTACGGGCGATAGACCCACAGGGCGAG 4344
QY 1469 GlnHisMetTyrArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaAlaIleLeu 1488
DDB 4345 CAGCAAAATGTGCGAGGCAATCCGCGAGTCTGTTAAAAACACAGAGAGAGGTGCTCTCTG 4404
```

```
QY 1489 ThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSer 1508
DDB 4405 ACCACCCATAACCTGGCTGAGCGGAGCCTTGTGTGACCGTGTGCCCATCATGTGTCT 4464
QY 1509 GlyGlnLeuArgCysIleLeuThrValGlnHisLeuLysSerLysPheGlyLysGlyTyr 1528
DDB 4465 GGAAGGCTTAGATGATGGCTCCACACCTGAAAAACAGACTTGGCAAGGATTAC 4524
QY 1529 PheLeuGluIleLysLeuLysAspTyrPheGluAsnLeuGluValAspArgLeuGlnArg 1548
DDB 4525 ATTCTAGAGCTAAAAAGTGAAG-----GAAAGCTCTCAAGTGACTTTGGTCCACACT 4575
QY 1549 GluIleGlnTyrIlePheProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAla 1568
DDB 4576 GAGATTCTGAAGCTTTTCCACAGGCTGCAGGCGAGGAAGGTATTCTCTTTGTTAAC 4635
QY 1569 TyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGlu 1588
DDB 4636 TATAAGTCGCCGTGGCAGACGCTTTACCTCTATCACAGACCTTTACAAATTAGAGCA 4695
QY 1589 AlaLysHisAlaPheAlaIleGluGlnTyrSerPheSerGlnAlaThrLeuGlnVal 1608
DDB 4696 GTGAAGCATTAACCTTTAACTGGAGAAATACAGCCTTTCTCAGTGACACTGGAGAGGTA 4755
QY 1609 PheValGluLeuThrLysGluGlnGluGluAspAsnSerCysGlyThrLeuAsnSer 1628
DDB 4756 TTCTTAGAGCTTTCTAAAGAACAG---GAAGTAGGAATTTTGTAGGAAGAAATTGATACA 4812
QY 1629 ThrLeuTyrTrpGlu 1633
DDB 4813 ACAATGAGATGAAA 4827
```

Search completed: April 13, 2004, 04:35:26

Job time : 1339 secs



Db 1234 CTATGGCAATATATCTTGCACAAATTTTACC-----TATGGAGATGAGCGCCATTAT 1287  
Qy 430 SerSerLeuTyrPheLeuLysProSerTyrTrpSerLysSerLysArgAsnTyrGluGlu 449  
Db 1288 TCTCCTTTATTTTCTTGAATTCATCTTGTTCACCAACCAAGGACATATGCTAAG 1347  
Qy 450 LeuSerGluGlyAenValAsnGlyAsnLysSerPheSerGluLeuLeuGluProValSer 469  
Db 1348 GTTATTGAGAAGAAATCGATCGTGAGCATCCCTCTGATGATATTTTGAACCGTAGCT 1407  
Qy 470 SerGluPheValGlyLysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLys 489  
Db 1408 CCGTAATTCGAAGGAAAGAACCATCAGATCAGAAATGTAAAGAGGAATATAAGGA 1467  
Qy 490 LysGlyGluAenValGluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIle 509  
Db 1468 AAATCTGGAAGTGAAGCAATTGAAGGCTTGCTCTTGCATATATGAGGTCAATC 1527  
Qy 510 ThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGly 529  
Db 1528 ACGCAATCTCGGTTCACAGTGGAGCTGGCAATCTTCACTCTAAATATTTTAAATGGA 1587  
Qy 530 LeuCysProSerAspGlyPheAlaSerIleTyrGlyHisArgValSerGluLeuAsp 549  
Db 1588 TTGCTCTGCACAGAGAGGATCAGTTACCATCTATATAAATCTCTCTGAATGCCAA 1647  
Qy 550 GluMetPheGluAlaArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAsp 569  
Db 1648 GACTTGGAGGAATCAGAAAGATTAACCTGGCTGTCTCAATTCATTTCAATTTGAC 1707  
Qy 570 ValLeuThrValGluGluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsn 589  
Db 1708 ATACTCACCGTGAAGGAAACCTCAGCTGTGTGCTAAATAAAGGATTCATCTAAAG 1767  
Qy 590 AsnIleLeuGlnGluValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLysAsp 609  
Db 1768 GAAGTGGACAGAGGTACACGAATATTATTGGAATTTGGACATGCCAAACATTCAGAT 1827  
Qy 610 AsnGlnAlaLysLysLeuSerGlyGlnLysArgLysLeuSerLeuGlyIleAlaVal 629  
Db 1828 AACCTTGTCTAAACATTTAAGTGAAGGACAGAAAGAGCTGACTTTTGGGATTAACCAT 1887  
Qy 630 LeuGlyAsnProLysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSer 649  
Db 1888 TTAGAGATCCTCAATTTTGTCTTTAGTAGAACCACTACTGGATTGGATCCCTTTTC 1947  
Qy 650 ArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSer 669  
Db 1948 AGAGATCAAGTGTGGAGCCTCTGAGAGCGTAGACAGATCATGTGATCCTTTTCAGT 2007  
Qy 670 ThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGly 689  
Db 2008 ACCAGTCCATGGATGAGGCTGACATCTGGCTGTATAGAAAGTCAATCATCTCCAAATGG 2067  
Qy 690 MetLeuLysCysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrArg 709  
Db 2068 AGACTGAAGTGTGCAGGTTCTTCTATGTTTTTGAAGAGAGGCTGCTTGGATATCAC 2127  
Qy 710 LeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGln 729  
Db 2128 CTAAGTTACATAGGATGAAATATGTAAACCCAGAACAAATACATCTCTTACTACTCAT 2187  
Qy 730 HisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuPro 749  
Db 2188 CACATCCCCGATGCTAAATTAATAAACAAGAAAAAGCTTGTATATACATTTTGCCA 2247  
Qy 750 PheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeu 769  
Db 2248 CTGGGAAGGACAAATACATTTCCAGATCTTTTCAGTGATCTGGATTAAGTGTCTGCCAG 2307  
Qy 770 GlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGlu 789

Db 2308 GGAGTGACAGGTTATGACATTTCCATGTCAACTTAAATGAAGTCTTTTATGAAGTGGAA 2367  
Qy 790 ValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGlu 809  
Db 2368 GGACAGTCACTATCGAACAA--GATTTT-----GAACAAGTGGAGATG 2409  
Qy 810 GluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuIleLeuSerGluThr 829  
Db 2410 ATAAGAGACTCAGAAAGCTCAATGAATGGAGCTGGCTCACTCTCTCTCTGAAATG 2469  
Qy 830 LysAlaSerLeuValSerThrMetSerLeuTrpLysGlnMetTyrThrIleAlaLys 849  
Db 2470 CAGACAGCT--GTGAGTGACATGGGCTCTGGAGATGCAAGTCTTTGCCATGGCAGG 2526  
Qy 850 PheHisPhePheThrLeuLysArgLysSerValArgSerValLeuLeuLeuLeu 869  
Db 2527 CTCGCTTTCTTAAAGTAAACGTCACAACTAAA-----GTCTTATTGACCTTA 2574  
Qy 870 LeuIlePhePheThrValGlnIlePheMetPheLeuValHis-----HisSerPhe 886  
Db 2575 TTATTGGTATTTGGAAATCCCAATATTCCTTTGATGTTTGAATATATATATATGCTATG 2634  
Qy 887 LysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLysProGlyAsp 906  
Db 2635 TTAAATGAAAGATCGATTTGGGAATTTAAAAACGAATGTATTTCTCTCTCTGGACAA 2694  
Qy 907 LysProHisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAspSerAspIleSer 926  
Db 2695 CTTCCCGAAGAACCCCGTACCAGCTGTGTGATCATCAATAACAGAAATCAAAATTTGAA 2754  
Qy 927 AspLeuIleSerPhePheThrSerGlnAsnIleMetValThrMet----- 941  
Db 2755 GATTTTATAAATCACTGAAGCATCAAAATATACITTTTGGAGTAGATGACTTTGAAAC 2814  
Qy 942 IleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMetHis 961  
Db 2815 AGAAATGGTACTGATGGCCCTCTCA-----TACAATGGAGCTATCATATTTCTGTGT 2865  
Qy 962 SerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeuPro 981  
Db 2866 AAACAAGAGATATAGATTTTCAGTTGTGTATATACCAAGAGATTCGACTGTTTCCCA 2925  
Qy 982 IleLeuValAsnIleLeuSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIle 1001  
Db 2926 ATTCTTATGAATATATATCAGCAATGGCTACTCAAAATGTTTAAATCACACACACATAT 2985  
Qy 1002 GlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGluLeu 1021  
Db 2986 CGAATTGAGTCAAGCCCATTT-----CCTCTAGCCACATAGGACTCTGGACTGGGTG 3039  
Qy 1022 TyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAlaMet 1041  
Db 3040 CCGGATGGTTCCTTTTCTTCTTATTTTGGTTCTATGTAGCATTTCTCTTATATCACCATG 3099  
Qy 1042 GluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeu 1061  
Db 3100 GGCGATCAGTGAATACAGAAATAATGTAAGTCCAGCTATGGATTCAGGCTCTAC 3159  
Qy 1062 ProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleLeuLeu 1081  
Db 3160 ACTTCTGCTTACTGTGTGGGAGGACATGATGGAGCTCAGCTTCTTCTTCTTAAATCTTC 3219  
Qy 1082 IleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLys 1101  
Db 3220 CTTTAAATG-----TATTTAAATTTCTACATAGAAACATGCAGTACCTCTCTATT 3270  
Qy 1102 PheLeuAlaValValPheCysLeuIle-----GlyTyrValProSerValIle 1117  
Db 3271 ACAAGCCAAATCTGTTTGTCTTGGTTATAGTTACTCTCTGGTTATGCAGCTTCTCTGTC 3330  
Qy 1118 LeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrp 1137  
Db 3331 TTCTTCATATATATATCATTTATTTTTCGCAAAAGGAGAAAAACAGTGGCCTTTGG 3390



Qy 790 ValGluAlaGluLeuAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGlu 809  
 Db 2543 GCACAGTCACTATCGAACAA---GATTC-----GAACAAGTGGAGATG 2584  
 Qy 810 GluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuLeuLeuSerGluThr 829  
 Db 2585 ATAAGAGACTCGAAAGCTCAATGAATGGAGCTGGCTCACTTCCTCTCTCGAAATG 2644  
 Qy 830 LysAlaSerLeuValSerThrMetSerLeuTyrLysGlnGlnMetTyrThrIleAlaLys 849  
 Db 2645 CAGACAGCT---GTGAGTGACATGGCCCTCTGAGAAATGCAAGTCTTGGCCATGCGCAGG 2701  
 Qy 850 PheHisPhePheThrLeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeu 869  
 Db 2702 CTCGGTTCCTTAAAGTTAAACGCTCAAACTAAA-----GTGTTATTGACCCCTA 2749  
 Qy 870 LeuLeuPhePheThrValGlnIlePheMetPheLeuValHis-----HisSerPhe 886  
 Db 2750 TTATCGGTATTGGAAATCCCAATATTCCTTTGATTGTGAAATATAATATATGCTATG 2809  
 Qy 887 LysAsnAlaValProLeuLysLeuValProAspLeuTyrPheLeuLysProGlyAsp 906  
 Db 2810 TTAATGAAAGATCATGTTGGGAATTTAAAAACGAATTTGATATTTCTCTCTCTGACAA 2869  
 Qy 907 LysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAspIleSer 926  
 Db 2870 CTTCCCGAGAACCCCGTACCGCTGTTGATCATCAATACACAGAAATCAATATTGAA 2929  
 Qy 927 AspLeuLeuSerPhePheThrSerGlnAsnIleMetValThrMet----- 941  
 Db 2930 GATTTTATAAAATCACTGAAGCATCAAAATATATATCTTTTGGAAATGATGACTTTGAAAC 2989  
 Qy 942 IleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMetHis 961  
 Db 2990 AGAATGGTACTATGCGCTCTCA-----TACATGGAGCTATCATAGTTTCTGCT 3040  
 Qy 962 SerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeuPro 981  
 Db 3041 AAACAAAGAGATATAGATTTTTCAGTTGTGTGTAATACCAAGAGATGCACTGTTTCCA 3100  
 Qy 982 IleLeuValAsnIleLeuSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIle 1001  
 Db 3101 ATTCTTATGATATATATACAGATGGCTACTTCAAAATGTTTAAATCACACACATAT 3160  
 Qy 1002 GlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGluLeu 1021  
 Db 3161 CGAATTGAGTCAAGCCCATTT-----CCTCTTAGCCACATAGGACTCTGGACTGGT 3214  
 Qy 1022 TyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAlaMet 1041  
 Db 3215 CCGGATGGTTCCTTTTCTTATTTTGGTCTATGATGATGATTTCTCTTATATCACCATG 3274  
 Qy 1042 GluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeu 1061  
 Db 3275 GGCAGCATCAGTATACAGAAATATGTAAGTCCAGCTATGATTTTCAGCCCTCTAC 3334  
 Qy 1062 ProSerAlaTyrTrpIleGlnAlaValAspIleProLeuPhePheIleLeuLeu 1081  
 Db 3335 ACTTCTGCTACTGTTGGTGGCAGGCACTAGTGGAGCTCAGCTTCTTCAATTTTAAATCTC 3394  
 Qy 1082 IleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLys 1101  
 Db 3395 CTTTAAATG-----TATTTAATTTCTACATAGAAACATGCACTACCTCTTATT 3445  
 Qy 1102 PheLeuAlaValValPheCysLeuLeu-----GlyTyrValProSerValIle 1117  
 Db 3446 ACAAGCCAAATTTGTTGTTGTTGTTATAGTTACTCTCTGTTTATGACCTTCTCTGTC 3505  
 Qy 1118 LeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuLeuLeuLeuLeuThrTrp 1137  
 Db 3506 TTCTTAT 3565

Qy 1138 SerPheLeuTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePhe 1157  
 Db 3566 TCATTTTACTTCTTTTTCCTCCACCATCATGTTTTCATCCTTAAATCAATCATTTT 3625  
 Qy 1158 ---MetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIle 1176  
 Db 3626 GACTAAGTATATATGATTACCACATG-----GTATTGTTCTCTTCA 3667  
 Qy 1177 TyrProLeuLeuGlyCysLeuLeuSerPheIleLysIleSerTrpLysAsnValArgLys 1196  
 Db 3668 TATACCTTGTCTGGA---TTTAAACCTTTTGGAAAGTG-----AGA 3706  
 Qy 1197 AsnValAspThrTyrAsnProTyrAspArgLeuSer----- 1208  
 Db 3707 GACGAGAGACATACAGAGAAATTTCCAGAGGCAATTTGAATTGAGTCCCACTGATTTT 3766  
 Qy 1209 ValAlaValIleSerProTyrLeuGlnCysValLeuTyrIlePheLeuLeuGlnTyrTyr 1228  
 Db 3767 CTAGTCTGCTTCATACCTACTTTCAGACTTTCCTATTCGTTTCTTCTTCAAGATCATG 3826  
 Qy 1229 GluLysLysTyrGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThr 1248  
 Db 3827 GAATAAATGCTGAAAGAAAGAAATGCGAAAGATCCTGTTTTCAGA---ATTTCCTCC 3883  
 Qy 1249 LysSerLysAsnArgLysLeuProGluProProAspAsnGluAspGluAspVal 1268  
 Db 3884 CAAAGTAGAGATGCTAAG---CCAAATCCAGAAAGAACCCATAGATGAAGATAT 3940  
 Qy 1269 LysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluLysValPro 1288  
 Db 3941 CAAACAGAAAGATAGAACAGCCTGCTCTGACCATCTTCAATCTTAGATGAGAAACCT 4000  
 Qy 1289 SerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSer 1308  
 Db 4001 GTTATATTCCAGCTGCTCTACACAAAGAAATATGACGGCCAGAGAAAGTTGCTTTTCA 4060  
 Qy 1309 ArgLysValLysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIle 1328  
 Db 4061 AAGAGGAAGAAAGAAATAGCAGCAAGAAATATCTTCTGTTGTTTCAAGAGGTGAAT 4120  
 Qy 1329 LeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThrIleLeuAsnIleLeuValGly 1348  
 Db 4121 TTGGGATTGTAGACCCAGCTGCTGCTGGAAGAAAGTTTCAATCTTATAGATGATCTGGG 4180  
 Qy 1349 AspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGlu 1368  
 Db 4181 ATCACAAGGCCCACTGCTGAGAGGTGGAACCTGAAAGCGTGCAGTTCA----- 4228  
 Qy 1369 AspAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTyrProAsp 1388  
 Db 4229 -----GTTTGGCCACCTGGGGTACTGCGCTCAAGAGAACCTGCTGTGCCCATG 4279  
 Qy 1389 ThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAsp 1408  
 Db 4280 CTGACGTTGAGGGAACACCTGGAGGTGATGCTGCGCTCAAGGGGCTCAGGAAAGCGGAC 4339  
 Qy 1409 MetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLys 1428  
 Db 4340 GCGAGGCTCCCATCGCAAGATTAGTACGTCCTTCAACTGCATGACGAGCTGAATGTT 4399  
 Qy 1429 ThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeu 1448  
 Db 4400 CCTGTGCAGAAATTAACAGCAGCAATCAGCAGAAAGTTGTTTGTGCTGAGCCCTCTG 4459  
 Qy 1449 GlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLys 1468  
 Db 4460 GGAACATCAGCTGTTGCTTCTCGATGAACCATCTACGGGCATAGACCCACAGGCGCAG 4519  
 Qy 1469 GlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaIleLeu 1488  
 Db 4520 CAGCAATGTGGCAGGCAATCCAGGCACTGTTTAAAAACACAGAGAGAGGTGCTCTCTG 4579  
 Qy 1489 ThrThrThrIleTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSer 1508

362	Db	362	AATCTGGGAAGGGTAGATAAAATTAATAGCTCTTCTTTAAATGGTGTGTATACACCAATA	421
85	Qy	85	ThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu-----Pro	101
422	Db	422	TCTAAATTAAACCCAGCAGATAATGAATAAAACAGACACTTGCTCCTCTTTGAAAGGAACA	481
102	Qy	102	AspValIleIleThrGluGluTyThrAsnGlnLysGluMetLeuThrSerSerLeuSer	121
482	Db	482	AGTGTCATTTGGGGCACCACCAATAAACACACATGACGACGAATATCTTCGGAATAATTACCA	541
122	Qy	122	LysProSerAsnPheValGlyValPheLysAspSerMetSerTyrGluLeuArgPhe	141
542	Db	542	TATGCT-----ATGGGAATCATCTTTAATGAAACTTCTCTTATAGTTAATATTT	592
142	Qy	142	PheProAspMetIleProValSerSerIleTyrMet---AspSerArgAlaGlyCys---	159
593	Db	593	TTCCAGGGGATAT-----AACAGTCCCATTTTGGAAAGAAGATTCTCAGCTCATTTGCTGG	646
160	Qy	160	-----SerLysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrVal	175
647	Db	647	GATGGATATGCTGAGTTTTCATGTACATTGACCAAAATCTGGAAATAGAGGATTTGTGGCT	706
176	Qy	176	LeuGlnAlaSerIleAspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLys	195
707	Db	707	TTCAAAACAGCATTAATACTGCCATATAGAAATCAACAACCATCACCTGCTGATGGAG	766
196	Qy	196	GluLeuGluSerThrLysAlaValIleMetGlyGluThrAlaValGluIleAspThr	215
767	Db	767	GAGTTGATGTCAGTTACTGCTATTAACATATGAAGACATTTACCTTCACTAACTAAAAATCTT	826
216	Qy	216	PheProArgGlyValIleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPhe	235
827	Db	827	CTTCACAATGAGATGTTTATTTTATCTCTGCTTCATTTCTCCACCTGTGATATTTT	886
236	Qy	236	LeuAlaIleHisIleValAlaGluLysGluLysIleLysGluPheLeuLysIleMet	255
887	Db	887	ATATCACTCAATGTAACAAAAGAGAGA---AAAAAGCTCAAGAAATTTGATCAAAATGATG	943
256	Qy	256	GlyLeuHisAspThrAlaPheTyrPleuSerTyrPValLeuLeuTyrThrSerLeuIlePhe	275
944	Db	944	GCTCTCCAGATTTCAGCATCTGGCTCTCTGGGGTCTTAATCTATGCTGGCTTCATCTTT	1003
276	Qy	276	LeuMetSerLeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSer	295
1004	Db	1004	ATTATTTCCATATTCATTACAATATATCAATATCAACCAATATAGTCATGACTGGC	1063
296	Qy	296	SerIleValIlePheLeuLeuPheLeuTyrGlyLeuSerSerValPhePheAlaLeu	315
1064	Db	1064	TTCAATGGTCATATTTATCTCTTTTTTATGGCTTATCTTTGGTAGCTTTGGTGTTC	1123
316	Qy	316	MetLeuThrProLeuPheLysSerLysHisValGlyIleValGluPhePheValThr	335
1124	Db	1124	CTGATGAGTGTGCTTTAAAGAAAGCTGCTCCTACCAATTTGGTTGTGTTTCTCCTTACC	1183
336	Qy	336	ValAlaPheGlyPheIleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeu	355
1184	Db	1184	CTCTTTTGGGATGCTCTGGGA---TTCACGTATTTTATGTAACCAACTTCCTTCATCTCTG	1240
356	Qy	356	ValTrpLeuPhe-----SerProPheCysHisCysThrPheValIleGlyIleAla	372
1241	Db	1241	GAGTGGATTGTAATTTGCTAGCCCTTT-----GCCTTTACTCTGGATGATT	1291
373	Qy	373	GlnValMetHisLeuGluAspPheAsn---GluGlyAlaSerPheSerAsnLeuThrAla	391
1292	Db	1292	CAGATTATCAAACTG---GATTATAACTTGAATGGTGTAAATTTTCTGACCCCTTCAGGA	1348
392	Qy	392	GlyProTyrProLeuIleIleThrIleLeuMetLeuThrLeuAsnSerIlePheTyrVal	411
1349	Db	1349	GACTCATATACAATGATAGCAACTTTTCTATGTTGCTTTTGGATGCTCATCTACTTG	1408
412	Qy	412	LeuLeuAlaValTyrLeuAspGlnValIleProGlyGluPheGlyLeuArgArg-----	429

Db 3642 GGCACCATGTTA-----ATACCTCCCTTACATTGATGGCTCT 3680  
 Qy 1183 LeuileSerPheilleLysileSer-----TrpLysAsnValArgLysAsn 1197  
 Db 3691 CTATTCATTTTCTGAGATTCTCCGATTCATGATTCATGAGCTTACAGATCT 3740  
 Qy 1198 ValAspThrTyAsnProTrpAspArgLeuSerValAlaValileSerProTyLeuGln 1217  
 Db 3741 GAAATTGATATC-----CTGGCACTGCTAATACCTTACCTTCAT 3779  
 Qy 1218 CysValLeuTrpIlePheLeuLeuGlnTyTyTyGluLysLysTyGlyGlyArgSerIle 1237  
 Db 3780 TTTCTCATTTTCTTCTTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 3839  
 Qy 1238 ArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGlu 1257  
 Db 3840 AGAAGGATCTGCTTCTCAGA--ATTCTCCAGAGC---AACGCTATTTTCCAAAC 3893  
 Qy 1258 ProProAspAsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGlu 1277  
 Db 3894 CCAGAAGAGCCTGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3953  
 Qy 1278 LeuMetGlyCysGlnCysCysGluLysLysProSerIleMetValSerAsnLeuHysLys 1297  
 Db 3954 GCTATGGCTGCGAGACTTTGATGAGACACCCGTCATTCATGCTGCTACGAGAG 4013  
 Qy 1298 GluTyAspAspLysLysPheLeuLeuSerArgLysValLysLysValAlaThrLys 1317  
 Db 4014 GAATATGAGGCAAAAGAAATTTGCTTTCTTAAGAGGAAAGAAACAAATTCGCACAGA 4073  
 Qy 1318 TyriLeSerPheCysValLysLysGluLysLeuGlyLeuLeuLeuProAsnGlyVala 1337  
 Db 4074 AATGCTCTTTTGTGTTAAAGAGTGAAGTATAGGACTGTAGGACAAATGGAGCT 4133  
 Qy 1338 GlyLysSerThrIleAsnLeuValGlyAspIleGluProThrSerGlyGlnVal 1357  
 Db 4134 CGTAAAGTACAACTATTAAAGTATACTGAGACACAAACCACTGCGAGGACAGTG 4193  
 Qy 1358 PheLeuGlyAspTySerSerGluThrSerGluAspAspAspSerLeuLysCysMetGly 1377  
 Db 4194 ATTTTG-----AAAGGAGCGGTGGAGGGGACCCCTGGGCTTCTGGGG 4238  
 Qy 1378 TyCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGluile 1397  
 Db 4239 TACTGCTCAGGAGATCGCTGTGGCCCACTGACAGTGGAGGACCACTGGAGGTG 4298  
 Qy 1398 TyGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThr 1417  
 Db 4299 TAGCTGCGTGAAGGCTCTAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4358  
 Qy 1418 HisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyile 1437  
 Db 4359 GATGCGCTCAGCTGAGGACGAGCTGAAGGCTCCCGTGAAGACCTTGTACAGGGAATA 4418  
 Qy 1438 LysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAsp 1457  
 Db 4419 AAGCGAAAGCTCGCTTTGTGCTGAGCATCTGGGGGAAACCCGTCAGTGGTCTCTGGAT 4478  
 Qy 1458 GluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaileArgThr 1477  
 Db 4479 GAGCGCTGACCGGATGACCCCGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4538  
 Qy 1478 AlaPheLysAsnArgLysArgAlaIleLeuThrThrHisTyMetGluGluAlaGlu 1497  
 Db 4539 ACCTTTAGAAACACGAGGAGGGGCGCCCTCTGACCACTACATGCGAGAGGCTGAG 4598  
 Qy 1498 AlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrVal 1517  
 Db 4599 GCGGTGTGTGACCGAGTGCCCATCATGTGTGTCAGGAGGCTGAGATGTTGTTCCATC 4658  
 Qy 1518 GlnHisLeuLysSerLysPheGlyLysGlyTyTyPheLeuGluLysLeuLysAspTrp 1537  
 Db 4659 CAACACCTGAAAGCAAAATTTGGCAAGACACTCTGCTGGAGATGAGCTGAAG----- 4712

Qy 1538 IleGluAsnLeu---GluValAspArgLeuGlnArgGluileGlnTyTyIlePheProAsn 1556  
 Db 4713 -----AACCTGGCACAAATGGAGCCCTCCATGTCAGAGATCTCTGAGGCTTTTCCCCAG 4766  
 Qy 1557 AlaSerArgGlnGluSerPheSerSerIleLeuAlaTyLysIleProLysGluAspVal 1576  
 Db 4767 GCTGCTCAGCAGAAAGGTTCTCCCTGATGCTATAGTTCCTGTTGAGGATG 4826  
 Qy 1577 GlnSerLeuSerGlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGlu 1596  
 Db 4827 CCACCTTATCATCAGGCTTCTTCAATTAGAGATAGTAAACAGAGTTTCGACCTGAG 4886  
 Qy 1597 GluTySerPheSerGlnAlaThrLeuGlnValPheValGluLeuThrLysGluGln 1616  
 Db 4887 GAGTACAGCTCTCAGCTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4946  
 Qy 1617 Glu-----GluGluAsp 1620  
 Db 4947 GAGCTGGTGTATCTTGAAGAGGAC 4970  
 RESULT 14  
 US-10-005-338B-2  
 ; Sequence 2, Application US/10005338B  
 ; Publication No. US20030044895A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: DENELE, Patrice  
 ; APPLICANT: ROSIER-MONTUS, Marie-Francoise  
 ; APPLICANT: PRADES, Catherine  
 ; APPLICANT: ARNOULD-REQUIGNE, Isabelle  
 ; APPLICANT: DUVERGER, Nicolas  
 ; APPLICANT: ALLIKWETS, Rando  
 ; APPLICANT: DEAN, Michael  
 ; TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENE  
 ; FILE REFERENCE: ABCA5, 6, 9, 10  
 ; CURRENT APPLICATION NUMBER: US/10/005,338B  
 ; PRIOR FILING DATE: 2001-12-07  
 ; PRIOR APPLICATION NUMBER: US 60/263,231  
 ; PRIOR FILING DATE: 2001-01-23  
 ; PRIOR APPLICATION NUMBER: FR 00403440.1  
 ; PRIOR FILING DATE: 2000-12-07  
 ; NUMBER OF SEQ ID NOS: 217  
 ; SOFTWARE: Patent in Ver. 2.1  
 ; SEQ ID NO 2  
 ; LENGTH: 5296  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-005-338B-2  
 Alignment Scores:  
 Pred. No.: 9,46e-246 Length: 5296  
 Score: 3124.00 Matches: 688  
 Percent Similarity: 61.14% Conservative: 330  
 Best Local Similarity: 41.32% Mismatches: 553  
 Query Match: 37.08% Indels: 94  
 DB: 14 Gaps: 33  
 US-10-090-458-5 (1-1642) x US-10-005-338B-2 (1-5296)  
 Qy 5 IleArgLysValGlyValTrpArgGlnThrArgThrLeuLeuLysAsnTyriLeuile 24  
 Db 182 ATGAACAGAAAGAGCTGTATCAGCAACCAAGACATCTCTGCAAGAAATTTCTTAAG 241  
 Qy 25 LysCysArgThrLysLysSerValGlnGluLeuPheProLeuPhePheLeuPhe 44  
 Db 242 AATGAGGATGAAAGAGAGAGGCTTATTGGAATGGGCTCTCAATATCTTAGACTG 301  
 Qy 45 TrpLeuileLeuileSerMetMetHisProAsnLysLysTyTyGluGluValProAsnile 64  
 Db 302 TGATTGCTCTGTTTCCAGTCCATGAGAAATGTCAGTTTCTGGAATGCTCTCTAG 361  
 Qy 65 GluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLeuGlyTyTyThrProVal 84

QY 455 VallanGlyAsnIleSerPheSerGluIleLeuProValSerSerGluPheValGly 474  
 Db 1515 ACAGATTCGTGATCCTACCACTAATGAGTGTGTTGAACCAAGTGTCTCCAGAAATCTGTGGG 1574  
 QY 475 LysGluAlaIleArgIleSerGlyIleGlnIleLysThrTyrArgLysLysGlyGluAsnVal 494  
 Db 1575 AAGGAGCCATCAGAAATCTTAAATAAGAAATATGTCAGGAAAGTGTGAGAGAGTA 1634  
 QY 495 GluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuGly 514  
 Db 1635 GAAGCTTTGAAGGTTGGTGTGTGACATATATGAAGCCAGATCAGTCCCTCTTGGT 1694  
 QY 515 HisSerGlyThrGlySerThrLeuMetAsnIleLeuCysGlyLeuCysProSer 534  
 Db 1695 CACAGTGGAGCTGGAATAACTACCTGTTAAACATCTAGTGGTGTGCTAGTTCCCAAC 1754  
 QY 535 AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
 Db 1755 TCAGTTCAGTCATGCTGTATTAATCACACATTTCAAGATGGCTGATATAGAAAAATATC 1814  
 QY 555 ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
 Db 1815 AGCAAGTTCAGTGGATTTGTCACCAATCCCAATGTCATTTGGATTTCTCAGTGTGAAA 1874  
 QY 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnAsnIleGlnGlu 594  
 Db 1875 GAAAAACCTCAGGCTGTTGCTAAAAATAAAGGAGATTTGCCACATGAAGTGGAGAAAGAG 1934  
 QY 595 ValGlnLysValLeuAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLys 614  
 Db 1935 GTACACAGTGTGACAGGAATAGAAATGGAATATTCAGACATCTTGTCTCAAAAC 1994  
 QY 615 LeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLys 634  
 Db 1995 TTAAGTGTGGACAAATAGGAACTAACTTTTGGGATTTGGCATTTTAGGAGATCCTCAA 2054  
 QY 635 IleLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleValTyr 654  
 Db 2055 GTTTTGTATTGGATGACCGACTGCTGGATTTGGATCTCTTTCAAGCCACCGAATATGG 2114  
 QY 655 AsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPheMetAsp 674  
 Db 2115 AATCTCTCAAAGAGGGGAAATCAGACAGAGTAATCTCTCTCAGCACCCAGTTTATAGAT 2174  
 QY 675 GluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLysCysVal 694  
 Db 2175 GAGGCTGACATCTGCGGACAGGAGGTTCATATCAATGGGAAGCTGAAGTGTCCA 2234  
 QY 695 GlySerSerMetPheLeuLysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIle 714  
 Db 2235 GGCTCTCTCTGTTCTTAAAGAAAGAAATGGGGCATAGGCTACCATTTAAGTTTGCATCTG 2294  
 QY 715 AspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAla 734  
 Db 2295 AATGAAGGTGTGATCCAGAGAGTATAACATCAGTGGTTAAGCAGCAGCATCTCTGATGCC 2354  
 QY 735 ThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAsp 754  
 Db 2355 AAATTCAGACGACCAAGTCAAGAAACTGTATATATTTTGCCTTGGAAAGGACAAAC 2414  
 QY 755 LysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyr 774  
 Db 2415 AAATTTCCAGAACTTTTACAGGGATCTTGTATAGATGTTCTTAAACCAAGGCATTTGAGGATTAT 2474  
 QY 775 GlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGluIle 794  
 Db 2475 GTGTGTTTCCATAACAACTTTGATGAGGTGTTTCTGAAATTAGAAGGAAATCAACTATT 2534  
 QY 795 AspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluGluMetAspSerLys 814  
 Db 2535 GATGAATCAGATATTGGAATTTGGGACAAATTAACAACTGTAGTGGGGGCAAAAGATATAGGA 2594  
 QY 815 SerPheAspGluMetGlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuVal 834

Db 2595 AGCCTGTTGAGCTGGAAACAAGTTTGTCTTCCAGGAAACAAGGAAAAACA---ATC 2651  
 QY 835 SerThrMetSerLeuTyrIleGlnIleMetTyrThrIleAlaLysPheHisPhePheThr 854  
 Db 2652 AGTGGCGTGGCCTCTGGAGGAGGAGTCTGTGCAATAGCAAAAGTTCGCTTCCTTAAG 2711  
 QY 855 LeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeuLeuIlePhePheThr 874  
 Db 2712 TTAAGAAAGAAAGAAAGAAAGCCGTGGAGCTATATATTGCTTTTGGTATTAGCTTATC 2771  
 QY 875 ValGlnIlePheMetPheLeuValHisSerPheLysAsnAlaValValProIleLys 894  
 Db 2772 CTTCAACTTTGGAACTATTTTACAGATCATATCAGAAAAGT---TACCCGTGGGAA 2828  
 QY 895 LeuValProAspLeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSer 914  
 Db 2829 CTGCTCCAAATACATACATCTCTCCAGGACAAACACACAGATCCTCTGACCCAT 2888  
 QY 915 LeuLeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
 Db 2889 TTACTGTCATCAATAGACAGGCTCAACCATTTGATTAACCTTTTACATTCAGTGGCGA 2948  
 QY 935 GlnAsnIleMetValThrMet-----IleAsnAspSerAspTyrValSer 949  
 Db 2949 CAGACATAGCTATAGAGTGGATGCTTGGAACTAGAAATGGCACAGATGACCCATCT 3008  
 QY 950 ValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAla 969  
 Db 3009 -----TACAAATGGTGTATCATTTGTGTCAGGTGATGAAGGATCAGAGATTTTCA 3059  
 QY 970 AlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIleLeuSerAsn 989  
 Db 3060 ATAGCATGTATACAAACGGCTGAATGCTTCTGCTCTCTGATGCTATTAGCAAT 3119  
 QY 990 TyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTyrSerThrProPhePhe 1009  
 Db 3120 GGACTACTTGGAAATTTTAAATTCGTGAGAACACATTCAGACTGACAGAGACATTTT 3179  
 QY 1010 GlnGluIleThrAspIleVal-----PheLysIleGluLeuTyrPheGlnAlaLeu 1027  
 Db 3180 GAAGAGCATATGATGATGATGATGGTATGATGATGATGATGATGATGATGATGATG 3239  
 QY 1028 LeuGlyIleIleValThrAlaMetProProTyrPheAlaMetGluAsnAlaGluAsnHis 1047  
 Db 3240 -----CGAGCCTTTTCACTCCATATGCAATGAGCAGCATTTGGTGACTAC 3287  
 QY 1048 LysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrIle 1067  
 Db 3288 AAAAAAAGCTCATTTCCAGCTACGATTTCCAGGCTTACCCCTTCTGATGATGATGAT 3347  
 QY 1068 GlyGlnAlaValValAspIleProLeuPhePheIleLeuIleLeuMetLeuGlySer 1087  
 Db 3348 GGCCAGCATCTGTGTGATGTTTCCCTGTACTTTTGTATCTCTCTGCTAATG-----CAA 3401  
 QY 1088 LeuLeuAlaPheHisTyrGly-----LeuTyrPheTyrThrValLysPheLeuAla 1104  
 Db 3402 ATAATGATATATATTTTAGCCAGAGAGATATATATATATATATATATATATATAT 3461  
 QY 1105 ValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSer 1124  
 Db 3462 CAATCTCTGTAGTATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 3521  
 QY 1125 PheThrPheLysLysIleLeuAsnThrLysGluPheThrPheThrPheThrPheThrPhe 1144  
 Db 3522 TTTATTTTTCGAATGGGAAAAAATAGGGCATTTTGGTCAATTTTCTTCTTAAATG 3581  
 QY 1145 AlaLeuAlaCysIleAlaIleThrGluIleThr-----PhePheMetGlyTyrThrIle 1162  
 Db 3582 GTATCTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3641  
 QY 1163 AlaThrIleLeuHisTyrAlaPheCysIleIleIleIleIleIleIleIleIleIleIle 1182

Db	4749	CTGAGGCTTTTCCCCCGAGCGCTCAGCAGGAAGGTTCTCTCTCCCTGATGGCTATAAG	4808
Qy	1571	IleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluGluAlaLys	1590
Db	4809	TTGCGCTGTTGAGGATGTGCGACCTTTATCACAGGCTTCTTCAAAATTAGAGATAGTTAAA	4868
Qy	1591	HisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheVal	1610
Db	4869	CAGAGTTTCAGCTCTGGAGGAGTACAGCCTCTCACAGTCTACCTCTGGACAGGTTTTCGTG	4928
Qy	1611	GluLeuThrLysGluGlnGlu-----GluGluAsp	1620
Db	4929	GAGCTCTCCAGGACGAGGAGCTGGGTGATCTTCAAGAGAC	4970

## RESULT 13

```

US-10-005-338B-3
; Sequence 3, Application US/10005338B
; Publication No. US20030044895A1
; GENERAL INFORMATION:
; APPLICANT: DENEFFLE, Patrice
; APPLICANT: ROSIER-MONTUS, Marie-Francoise
; APPLICANT: PRADES, Catherine
; APPLICANT: ARNOULD-REGUIGNE, Isabelle
; APPLICANT: DUVERGER, Nicolas
; APPLICANT: ALLIKMETS, Rando
; APPLICANT: DEAN, Michael
; TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENES
; TITLE OF INVENTION: CONTAINING SUCH NUCLEIC ACIDS, AND USES THEREOF

```

```

1 FILE REFERENCE: AC0437, 6129, 10
2
3 CURRENT APPLICATION NUMBER: US/10/005,338B
4
5 CURRENT FILING DATE: 2001-12-07
6
7 PRIOR APPLICATION NUMBER: US 60/263,231
8
9 PRIOR FILING DATE: 2001-01-23
10
11 PRIOR APPLICATION NUMBER: FR 00403440.1
12
13 PRIOR FILING DATE: 2000-12-07
14
15 NUMBER OF SEQ ID NOS: 217
16
17 SOFTWARE: PatentIn Ver. 2.1
18
19 SEQ ID NO 3
20
21 LENGTH: 5981
22
23 TYPE: DNA
24
25 ORGANISM: Homo sapiens
26
27 US-10-005-338B-3

```

Alignment Scores:	
Pred. No.:	1.05e-251
Score:	3197.50
Percent Similarity:	61.35%
Best Local Similarity:	41.87%
Query Match:	37.95%
DB:	14
	14
Length:	5981
Matches:	690
Conservative:	321
Mismatches:	562
Indels:	75
Gaps:	27

US-10-090-458-5 (1-1642) X US-10-005-338B-3 (1-5981)

Oy	6	ArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLysAsnTyrLeuIleLys	25
Dy	:	:::::	
Db	153	AGACCATGAGCGTGCGTCAGCAACAATGGGTCTCTCTCAGAAACTGCTCAAAAAA	212
Oy	26	CysArgThrLysLysSerValGlnGlnIleLeuPhePro---LeuPheLeuPhe	44
Dy	:	:::::	
Db	213	TGAGAATGAAGAACAGACACCTGTGTGAATGGCTCTTTTCATTCTTCGGTACTGTTT	272
Oy	45	TrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGluValProAsnIle	64
Dy	:	:::::	
Db	273	CTGTACCTATTTTCTCCCAATTATCATCAAGTTTCATGCACATCTCTCAAATGCTTCCAATG	332
Oy	65	GluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLeuGlyTyrThrProVal	84
Dy	:	:::::	
Db	333	GAICTGGGACGTGTAGATAGTTTTTAATCATCATTAATTATGTTATTTGCACTTGGCACCTGAA	392
Oy	85	ThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeuProAspValIle	104
Dy	:	:::::	
Db	393	TCCAAACTACCACCAAGATATGAACAAAGTGGTCTCAGCCCCATTCCTTAAAGGAAGA	452

Qy	105	IleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLysProSer	124
Db	453	ACAATCATGGGGGCGCTGATGAAAGACATGATGATGGATTGAACTATTCATA	512
Qy	125	AsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArgPhe---	143
Db	513	GACGCAGTGAGATCATCTTTACTGATACCTTCCTCCATCCATTGGAAGTTTCTTG	572
Qy	144	AspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCys-----	159
Db	573	CATAGAAATCCCAATGATGAAGACACAGAGACCATTCAGTCCACGTCAAGCAGTGAAT	632
Qy	160	---SerIysSerCysGluAlaAlaGlnTyrTyrSerSerGlyPheThrValLeuGlnAla	178
Db	633	GAAGAAATGAAGTGTGAAGGTTTCAGAGTTCTGGGAGAAAGCGTTGTAGCTTTTCAAGCT	692
Qy	179	SerIleAspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGlu	198
Db	693	GCCATTAACTGCTGCTATCATGAATAATCGCAACAATCAATTCAGTGATGGAACAGCTGATG	753
Qy	199	SerThrLysAlaValIleMetGlyGluThrAlaValValGluIleLeuPheThrProArg	218
Db	753	TCAGTTACTGGGTACATATGAAGATATTACCTTTGTTGCCCAAGAGAGAGTTGCCAAT	812
Qy	219	GlyValIleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIle	238
Db	813	GATTTTTCATTTCTTTTGCATATTCTTTTCTACATTTATATACATGATGATCAGTC	872
Qy	239	HisIleValAlaGluLysGluLysLysIleGlyGluPheLeuLysIleMetGlyLeuHis	258
Db	873	AATGTTACACAAGAAAGACAA---TACATTACGTCATTGATGACAAATGATGGGACCTCCGA	929
Qy	259	AspPheAlaPheTyrLeuSerTyrValLeuLeuTyrThrSerLeuIlePheLeuMetSer	278
Db	930	GAGTCAGCATCTGGCTTTCCTGGGGTTTCATGATGCTGGCTCACTCCATCATCATGCGC	989
Qy	279	LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleVal	298
Db	990	ACTTTAATGGCTTTATTGTAATAATTCGCAAAATTCGTCCTCGATCGTTGTTGTGATG	1049
Qy	299	IlePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThr	318
Db	1050	GTCTTCACCTCTTCTCCCTATGGCTGTCCTTTGTAATCACTTAGCTTTCCTGATGAGT	1109
Qy	319	ProLeuPheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPhe	338
Db	1110	GTGTGTATAAGAAACCTTTCTTACGGGCTTGCTGTGTCTTCTCCCTATTGTCTTTTGG	1169
Qy	339	GlyPheIleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTyr---	357
Db	1170	GGGATCTCGGATTCCTCCA---GCATTGTATACACATCTCTCGATTTTGGATGGACT	1226
Qy	358	-----LeuPheSerProPheCysHisCysThrPheValIleGlyIleAlaGlnValMet	375
Db	1227	TTGTGTCTTCTTAGCCCTTT-----GCCTTCACTGTGGGATGGCCACGCTTATA	1277
Qy	376	HisLeuGlu---AspPheAsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyr	394
Db	1278	CATTGGACATGATGTAATCTAATGCCACTTGGATCTCT---TCACAAATCCATAC	1334
Qy	395	ProLeuIleIleThrIleIleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAla	414
Db	1335	CTCATAATAGTACTCTTTTCATGTTGGTTTGTGACACCCCTTCCTGATTTGGTATTGACA	1394
Qy	415	ValTyrLeuAspGlnValIleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPhe	434
Db	1395	TTATATTTTGACAAAAATTTGCCCGCTGAATATGGACATCGATGTTCTCCCTGTTTTC	1454
Qy	435	LeuLysProSerTyrTyrSerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsn	454
Db	1455	CTGAATCCCTGTTTTTGGTTTTCACACCGAAGGGCTAATCATGTGGTCTCTTGAGAATGAA	1514



QY	1203	ProTrpAspArgLeuSer-----ValalaVal 121
DB	3702	CCTCCTGATTTCCAGGATTCCTTAGGAGCTTCAGAAATCTGAAATGTTGATACCTGGCCACTG 3761
QY	1212	IleSerProTyrLeuGlnCysValLeuTyrPheLeuLeuGlnTyrTyrGluLeuLys 1231
DB	3762	CTGATCCCTTACCTTCATTTTTCATTTTCTTTTTCATTTCTGCGATCGCTAGAAATGAAC 3821
QY	1232	TyrGlyGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLys 1251
DB	3822	TGCAGGAAGAAACAAATAGAGAAAGATTCCTGTGTTCAGA---ATTTCCTCAAGAAGC--- 3875
QY	1252	AsnArgLysLeuProGluProAspAsnGluAspGluAspGluAspValLysAlaGlu 1271
DB	3876	AACCTATTTTTCCAAACCCAGAGAGACCTGAGGAGGAGGAGAGATATCCAGATCGAA 3935
QY	1272	ArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGluLysProSerIleMet 1291
DB	3936	AGAATGAGAACACAGTGAATGCTATGGCTGTGCGAGACTTGTATGAGACACCCGTCATCATTT 3995
QY	1292	ValSerAsnLeuHisLysGluTyrAspAspLysValysAspPheLeuLeuSerArgLysVal 1311
DB	3996	GCCAGCTGCTACGGAAGGAATATGACGGCAAAAGAAAAATTGCTTTTCTAAAGGAAG 4055
QY	1312	LysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIleLeuGlyLeu 1331
DB	4056	AAAACAATTGCCACAGAAATGTCTCTTTTGTGTAAAAAGGTGAAGTTATAGGACTG 4115
QY	1332	LeuGlyProAsnGlyAlaGlyLysSerThrIleLeuValGlyAspIleGlu 1351
DB	4116	TTAGGACACAATGGGCTGTAAAGPACAACTATTAAAGATGATACTGGAGACACAAAA 4175
QY	1352	ProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAsp 1371
DB	4176	CCAACTGCAGGACAGGTGATTGTG-----AAAGGGAGCGGTGGAGGGGAA 4220
QY	1372	SerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTyrProAspThrThrLeu 1391
DB	4221	CCCCTGGGCTTCCTGGGGTACTGCCCTCAGAGAAATGCGCTGTGGGCCCAACCTGCACAGT 4280
QY	1392	GlnGlnHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGlu 1411
DB	4281	AGGCAGACCTGAGGTGACGTGCGTGAAGGTCTCAGGAAAGGGACGCAATGATC 4340
QY	1412	ValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLys 1431
DB	4341	GCCATCACACGGTTAGTGGATGCGCTCAAGCTGCAGGACCCAGCTGAAGGCTCCCGTGAAG 4400
QY	1432	LysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnPro 1451
DB	4401	ACTTGTGCAGAGGAATAAGCGAAGACTGTGCTTGTGTGACGATCTCTGGGAACCCG 4460
QY	1452	GlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMet 1471
DB	4461	TCAGTGTGTCTCTGGATGAGCGCTGCACCGGATGGACCCCGAGGGGCGAGCAAAATG 4520
QY	1472	TyrArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaIleLeuThrThrHis 1491
DB	4521	TGGCAGGTGATTCGGGCCACTTTTAGAAACACGAGAGGGGGCCCTCTCTGACCAACCCAC 4580
QY	1492	TyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeu 1511
DB	4581	TACATGCGACAGGCTGAGGGCGTGTGTGACCGAGATGGCCATCATGTGTGTGAGAAAGGCTG 4640
QY	1512	ArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGlu 1531
DB	4641	AGATGTATTGGTTCCTCCAAACACTGAAAGCAAAATTTGGCAAAAGACTACCTGCTGGAG 4700
QY	1532	IleLysLeuLysAspTrpIleGluAsnLeu---GluValAspArgLeuGlnArgGluIle 1550
DB	4701	ATGAGAGCTGAAG-----ACCTTGGCAAAATGAGGCCCTTCATGACGAGATC 4748
QY	1551	GlnTyrIlePheProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAlaTyrLys 1570



Db 3538 CTATTCAATTTTCTGAGATT----- 3558  
 QY 1203 ProTrpAspArgLeuSer-----ValalaVal 1211  
 Db 3559 COTCTGATTCCTAGGATTCCTTAGGAGCTTCAGAAATCTGAAATGTATATACCTGGCACTG 3618  
 QY 1212 IleSerProTyrLeuGlnCysValLeuTyrPheLeuLeuGlnTyrTyrGluLysLys 1231  
 Db 3619 CTGATCCCTTACCTTCATTTCTCATTTTCTTTCTATCTCGCATGCTAGAAATGAAAC 3678  
 QY 1232 TyrGlyGlyArgSerIleArgLysAspProPheArgAsnLeuSerThrLysSerLys 1251  
 Db 3679 TGCAGAGAAACTAATAGAGAAAGATCCTGTGTTCAGAA--ATTCTCTCAAGAGC--- 3732  
 QY 1252 AsnArgLysLeuProGluProProAspAsnGluAspGluAspGluAspValLysAlaGlu 1271  
 Db 3733 AACGCTATTTTCCAAACCCAGAGAGCTGAGAGAGAGAGAGATATCCAGATGAA 3792  
 QY 1272 ArgLeuLysValLysGluLeuMetGlyCysGlnCysGluGluLysProSerIleMet 1291  
 Db 3793 AGAATGAGACAGTGAATGCTATGCTGTGGAGACTTTGTATGAGACACCGCTCATCAT 3852  
 QY 1292 ValSerAsnLeuHisLeuGluTyrAspAspLysLysAspPheLeuLeuSerArgLysVal 1311  
 Db 3853 GCCACTGCTACGGAAGATATCGAGCAAAAGAAATGCTTTCTTAAAGGAAG 3912  
 QY 1312 LysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluLeuLeuGlyLeu 1331  
 Db 3913 AAAACAATTGCCACAGAATGCTCTTTTGTGTAAAGAGTGAAGTATAGGACTG 3972  
 QY 1332 LeuGlyProAsnGlyAlaGlyLysSerThrIleLeuLeuValGlyAspIleGlu 1351  
 Db 3973 TTAGACACAAATGGGCTGGTAAAGTACAACTATTAGATGATAATGGAGACACAAA 4032  
 QY 1352 ProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAsp 1371  
 Db 4033 CCNACTGCAGACAGTGAATTG-----AAAGGAGCGGTGGAGGGAA 4077  
 QY 1372 SerLeuLysCysMetGlyTyrCysProGlnLeuProLeuTyrProAspThrLeu 1391  
 Db 4078 CCCCTGGGCTTCCTGGGGTACTGCTCCAGGAGATGGCTGTGGCCCAACCTGCACAGTG 4137  
 QY 1392 GlnGluHisPheGluLeuTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGlu 1411  
 Db 4138 AGGAGACACTGGAGGTAGCTGCTGCTGAAAGTCTCAGAAAGGGAGCGCAATGATC 4197  
 QY 1412 ValLeuSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLys 1431  
 Db 4198 GCCATCACAGGTTAGTGGATCGCTCAAGCTGCAGGACCAAGCTGAAGGCTCCCGTGAAG 4257  
 QY 1432 LysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnPro 1451  
 Db 4258 ACCTTGTGAGAGGAGNATTAACGAAAGCTGTGCTTTGTGAGCATCCTGGGAAACCCG 4317  
 QY 1452 GlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMet 1471  
 Db 4318 TCAGTGGGTCTTCGATGAGCCCTCGACCGGATGACCCCGAGGGCGACGCAAAATG 4377  
 QY 1472 TrpArgAlaIleAspThrAlaPheLysAsnArgLysArgAlaIleLeuThrThrHis 1491  
 Db 4378 TGGCAGGTGATTCGGGCCACCTTTAGAAACACGAGAGGGCGGCCCTTCGACACCCAC 4437  
 QY 1492 TyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeu 1511  
 Db 4438 TACATGCGAGAGGCTGAGGCGGTGTGACCGAGTGGCCATCATGTGTACAGAGGCTG 4497  
 QY 1512 ArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGlu 1531  
 Db 4498 AGATGTTATGTTCTCCACACCTGAAAGCAAAATTTGGCAAGACTACCTGCTGGAG 4557  
 QY 1532 IleLysLeuLysAspTyrIleGluAsnLeu---GluValAspArgLeuGlnArgGluIle 1550  
 Db 4558 ATGAGCTGAAG-----AACCTGGCAAAATGAGGCCCTTCATGACAGATC 4605

QY 1551 GlnTyrIlePheProAsnAlaSerArgGlnGluSerPheSerSerIleLeuAlaTyrLys 1570  
 Db 4606 CTGAGGCTTTTCCCCAGGCTGCTCAGCAGGAAAGGTTCTCTCCCTGATGCTATAAG 4665  
 QY 1571 IleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGluAlaLys 1590  
 Db 4666 TTCCTGTGTGAGATGTGGACCTTTATCAGAGCTTTCTCAATTAGAGATAGTTAA 4725  
 QY 1591 HisAlaPheAlaIleGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheVal 1610  
 Db 4726 CAGAGTTTCGACCTGAGGAGTACAGCTCTCACAGTCTACCTGGAGCAGGTTTCTCTG 4785  
 QY 1611 GluLeuThrLysGluGlnGlu-----GluGluAsp 1620  
 Db 4786 GAGCTCTCCAAGAGCAGGAGCTGGGTGATCTTTGAAGAGGAC 4827

RESULT 12

US-10-090-454-1  
 ; Sequence 1, Application US/10090454  
 ; Publication No. US20020123106A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chen, Hongyun  
 ; APPLICANT: Le Bihan, Stephanie  
 ; APPLICANT: Nathwani, Farimal S.  
 ; APPLICANT: Connop, Bruce P.  
 ; TITLE OF INVENTION: NOVEL ABCA9 TRANSPORTER AND USES THEREOF  
 ; FILE REFERENCE: 100103 405  
 ; CURRENT APPLICATION NUMBER: US/10/090,454  
 ; CURRENT FILING DATE: 2002-03-04  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1  
 ; LENGTH: 5018  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-090-454-1

Alignment Scores:  
 Pred. No.: 4,56e-252 Length: 5018  
 Score: 3200.50 Matches: 690  
 Percent Similarity: 61.00% Conservative: 319  
 Best Local Similarity: 41.72% Mismatches: 558  
 Query Match: 37.98% Indels: 87  
 DB: 13 Gaps: 27

US-10-090-458-5 (1-1642) x US-10-090-454-1 (1-5018)

QY 6 ArgGluValGlyValTyrArgGlnThrArgThrLeuLeuLysAsnTyrLeuLys 25  
 Db 153 AGACGATGAGCTGGTGCAGCAACATGGGCTCTCTCGCAAGACTGCTCAAAA 212  
 QY 26 CysArgThrLysLysSerSerValGlnGluLeuLeuPhePro---LeuPhePheLeuPhe 44  
 Db 213 TGGAGAATGAAAGACAGACCTTGTGGAATGGCTCTTTTCATTTCTTCTGCTGTT 272  
 QY 45 TrpLeuLeuLeuLysSerMetMetHisProAsnLysLysTyrGluGluValProAsnIle 64  
 Db 273 CTGTACCTATTTTCTCAATTTTACATCAAGTTACATGACCTCCCTCAATGTTCTCAATG 332  
 QY 65 GluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLeuGlyTyrThrProVal 84  
 Db 333 GATCTGGGACGTAGATAGTTTAAATGATATACTAATTAATGTTATGTCATTGGCCTGAA 392  
 QY 85 ThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeuProAspValIle 104  
 Db 393 TCCAAACCTACCAAGAGATAATGAACAAAGTGGCTTCAGGCCCTTCTCAAAAGGAAGA 452  
 QY 105 IleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLysSerProSer 124  
 Db 453 ACAATCATGGGTGGCTGTGATGAAGAAAGCATGATGAATGGATTGGAATTTCAATA 512  
 QY 125 AsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArgPhe---PhePro 143

QY 475 LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnVal 494  
 Db 1432 AAAGAAGCCATCAGAAATCTTAAAGAAAGATATGAGGGAAGTGTGAGAGATA 1491  
 QY 495 GluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuGly 514  
 Db 1492 GAAGCTTTGAAGAGGTGTGGTGTTCACATATATGAAGGCCAGATCACTGCCTCTCTGGT 1551  
 QY 515 HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLysProSer 534  
 Db 1552 CACAGTGGAGCTGGAAAACCTCCCTGTTAAACATCTTAGTGGGTGTCTAGTTCACAA 1611  
 QY 535 AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
 Db 1612 TCAGGTTCACTGCTCTATATACACACATCTTCAAGAATGGCTGATATAGAAAATATC 1671  
 QY 555 ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
 Db 1672 AGCAAGTTCACGTGGATTTGTCCCAATCCCAATGTGCAATTTGGATTTCTCACTGTGAAA 1731  
 QY 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIlePheAlaAsnIleIleGlnGlu 594  
 Db 1732 GAAACCTCAGCTGTTGCTAAATATAAGGATTTGCCACATGAAGTGGAGAAGAG 1791  
 QY 595 ValGlnLysValLeuAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLys 614  
 Db 1792 GTACAACGAGTGTACAGGAATTAGAAATGGAATAATTTCAAGACATCTTGTCTCAAAAC 1851  
 QY 615 LeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLys 634  
 Db 1852 TTAAGTGTGGCAAAATAGGAAATTAACCTTTGGGATTTGCCATTTTAGGAGATCTTCAA 1911  
 QY 635 IleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleValTrp 654  
 Db 1912 GTTTTGTCTATTGATGCAACCGACTGCTGGATTTGGATCTCTTCAAGCAGCGAATATGG 1971  
 QY 655 AsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPheMetAsp 674  
 Db 1972 AATCTCTGAAAGAGGGAATACACAGAGTAATCTCTTCAGCACCAGTATTATAGAT 2031  
 QY 675 GluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLysCysVal 694  
 Db 2032 GAGGCTGACATCTGGCGACAGGAAGTGTTCATATCAATGGGAAGCTGAAGTGTCA 2091  
 QY 695 GlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrArgLysSerMetTyrIle 714  
 Db 2092 GGCTCTTCTCTTCTTAAAGAAAGAAATGGGGCATAGGCTACCAATTAAGTTTGCATCTG 2151  
 QY 715 AspLysTyrCysAlaThrGluSerLeuSerLeuValLysGlnHisIleProGlyAla 734  
 Db 2152 AATGAAGGTGTGATCCAGAGAGTATAACATCACTGGTTAAGCAGCATCTCTGATGCC 2211  
 QY 735 ThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAsp 754  
 Db 2212 AATGACAGCAAGAAAGTGAAGAAACCTTGTATATATTTGGCTTTGGAAAGAGCAAAAC 2271  
 QY 755 LysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyr 774  
 Db 2272 AATTTCCAGAACTTTACAGGATCTTGTATAGATCTTCAACCAAGGCAATGAGATAT 2331  
 QY 775 GlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGluIle 794  
 Db 2332 GGTGTTTCCATACCAACTTGAATGAGGTGTCTGAAATTAGAAGGAAATCAACATATT 2391  
 QY 795 AspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluMetAspSerLys 814  
 Db 2392 GATGAATCAGATATTGGAATTTGGGACATTTACAACTGATGGGGCAAAAGATATAGGA 2451  
 QY 815 SerPheAspGluMetGlnSerLeuLeuLeuLeuSerGluThrLysAlaSerLeuVal 834  
 Db 2452 AGCCTTGTGTGAGCTGCAAGCTTTTGTCTCTTCCACGAAACAAAGGAAACA---ATC 2508  
 QY 835 SerThrMetSerLeuTrpLysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThr 854

Db 2509 AGTGGCGTGGCGCTCTGGAGGAGCAGAGTCTCTGCAATAGCAAAAGTTCCTTAAG 2568  
 QY 855 LeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeuLeuLeuPhePheThr 874  
 Db 2569 TTAAGAAAGAAAGAAAGAAAGCTGTGGACTATATTATTTGCTTTTGGTATTAGCTTTATC 2628  
 QY 875 ValGlnIlePheMetPheLeuValHisSerPheLysAsnAlaValValProIleLys 894  
 Db 2629 CTTCACTTTTGGAAACATCTATTCTACGAGTCTATACAGAAAAGT---TACCGTGGGA 2685  
 QY 895 LeuValProAspLeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSer 914  
 Db 2686 CTGCTCCAAATACATACCTCTCTCACCAGACAAACACACAGATCTCTGACCCAT 2745  
 QY 915 LeuLeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
 Db 2746 TTAAGTGTCTATCATTAAGACAGGTCACCATTTGATATCTTTTACATTCACTGAGCGCA 2805  
 QY 935 GlnAsnIleMetValThrMet-----IleAsnAspSerAspTyrValSer 949  
 Db 2806 CAGAACATAGCTATAGAGTGGATGCTTTGGAACATAGAAATGGCAGATGACCCATCT 2865  
 QY 950 ValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAla 969  
 Db 2866 -----TACAATGGTGTCTATCATTTGTGAGTGTGAGGATGAAAGGATCAGAGATTTTCA 2916  
 QY 970 AlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIleSerAsn 989  
 Db 2917 ATAGCATGTAATACAAACCGGTGAATTCCTTCTGCTCTCTGATGTCTATAGCAAT 2976  
 QY 990 TyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTrpSerThrProPhePhe 1009  
 Db 2977 GGAATCTTGGAAATTTTAAATTCGTCAGAACACATTCAGACTGACAGAGACGACATTTT 3036  
 QY 1010 GlnGluIleThrAspIleVal-----PheLysIleGluLeuTyrPheGlnAlaLeu 1027  
 Db 3037 GAAGGCATATGATATGATGATGGTACCGAAGTAACACCTCTCTCTGATACCGATG 3096  
 QY 1028 LeuGlyIleValThrAlaMetProTyrPheAlaMetGluAsnAlaGluAsnHis 1047  
 Db 3097 -----GCAGCTCTTTTCACTCCATCATTTGCAATGAGCAGCATTTGGTACTAC 3144  
 QY 1048 LysIleValAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIle 1067  
 Db 3145 AAAAAAAGCTCATTTCCAGCTACGATTTTCCAGCTCTACCCCTCTGCACTGCTTT 3204  
 QY 1068 GlyGlnAlaValValAspIleProPhePheIleLeuLeuMetLeuGlySer 1087  
 Db 3205 GGCCAGCACTGGTGGATGTTTCCCTGTACTTTTGTATCTCTCTGCTCTGCTAATG-----CAA 3258  
 QY 1088 LeuLeuAlaPheHisTyrGly-----LeuTyrPheTyrThrValLysPheLeuAla 1104  
 Db 3259 ATATGATATATTTTATGCCAGGAGATATATTTATTAATTCAAAACCTGTTAAT 3318  
 QY 1105 ValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSer 1124  
 Db 3319 CAAATCCTGTGTAGTATTGGCTATGCTCATCTCTGTTTCTTGCATATGATGTTTCA 3378  
 QY 1125 PheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyrSerValAla 1144  
 Db 3379 TTTCAATTTTGGCAATGGGAGAAAAAATAGTGCAATTTGGTCAATTTTCTTCAATATG 3438  
 QY 1145 AlaLeuAlaCysIleAlaIleThrGluIleThr-----PhePheMetGlyTyrThrIle 1162  
 Db 3439 GTCATCTTCTCGATAGTGTCTACTGATCTAAATGAATATGATGATTTCTAGGCTATT 3498  
 QY 1163 AlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCys 1182  
 Db 3499 GGCACCATGTA-----ATACCTCCCTTCACTGATTTGGCTCT 3537  
 QY 1183 LeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsn 1202

```

Db 5230 CACACAGAGATTTCAGACTTTCCACAGAGCTGCTCGCAGGAAAGATATTCCTCTTTA 5289
Qy 1567 LeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeu 1586
Db 5290 ATGGCGTATAGTTACCTGTGAGGATGTCACACCTCTATCTCGGGCTTTTCAAGTTA 5349
Qy 1587 GluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThrLeuGlu 1606
Db 5350 GAGCGATGAACACAGACCTTCAACCTGGAGGAATACAGCCTCTCTCAGGCTACCTTGAG 5409
Qy 1607 GlnValPheValGluLeuThrLysGluGlnGluGluAspAsnSerCysGlyThrLeu 1626
Db 5410 CAGGTATTTCTAGACTCTGTAAAGACAG--GAGTGGGAATGTTGATGATAAAATT 5466
Qy 1627 AsnSerThrLeuTyrTrpGlu---ArgThrGlnGluAsp 1638
Db 5467 GATACACAGTTGAATGTAAGTAACTTCTCCACAGGAAGAC 5505

RESULT 11
US-10-090-454-3
; Sequence 3, Application US/10090454
; Publication No. US20020123106A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hongyun
; APPLICANT: Le Bihan, Stephane
; APPLICANT: Nathwani, Parimal S.
; APPLICANT: Connop, Bruce P.
; TITLE OF INVENTION: NOVEL ABCA9 TRANSPORTER AND USES THEREOF
; FILE REFERENCE: 100103.405
; CURRENT APPLICATION NUMBER: US/10/090.454
; CURRENT FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4875
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-090-454-3

Alignment Scores:
Pred. No.: 4,37e-252 Length: 4875
Score: 3200.50 Matches: 690
Percent Similarity: 61.00% Conservative: 319
Best Local Similarity: 41.72% Mismatches: 558
Query Match: 37.98% Indels: 87
Gaps: 27

US-10-090-458-5 (1-1642) x US-10-090-454-3 (1-4875)
Qy 6 ArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLysAsnTyrLeuLys 25
Db 10 AGACGCATGACGTGGGTGACAAACATGGGCTCTTCTGCAAGAACTGCTCAAAAAA 69
Qy 26 CysArgThrLysLysSerSerValGlnGluIleLeuPhePro---LeuPhePhePhe 44
Db 70 TGGAGATGAAGAACACAGACCTTGTGGAATGGCTCTTTTCATTTCTCTGGTACTGTT 129
Qy 45 TrpLeuLeuLeuSerMetMetHisProAsnLysLysTyrGluGluValProAsnIle 64
Db 130 CTGTACCTATTTTCTCAATTTATCATCAAGTTCATGACACTCTCTCAAAATGCTTCAATG 189
Qy 65 GluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGlyTyrThrProVal 84
Db 190 GATCTGGACGTGATAGTATTTAATGATCACTAATATGTTATGCTATTCACCTCGAA 249
Qy 85 ThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeuProAspValIle 104
Db 250 TCCAAAACCTCCCAAGAGATAATGAACAAAGTGCTTCAGGCCATTCCTAAAAGGAAGA 309
Qy 105 IleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLysProSer 124
Db 310 ACAATCATGGGGTGGCTGATGAAGAAAGCATGGATGAATTGGATTGAACTATTCATA 369

```

```

Qy 125 AsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArgPhe---PhePro 143
Db 370 GACGCAGTGAGAGTCACTTTTACTGATACCTCTCTCTACCATTTGAAGTTTCTTTGGGA 429
Qy 144 AspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCys----- 159
Db 430 CATAGAAATCCCATGATGAAGACACAGACCACTTCAGCTCAGCTGTCAAGCAGTGAAT 489
Qy 160 ---SerLysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAla 178
Db 490 GAAAAAATCAAGTGTGAGGTTTCAGAGTCTCGGAGAAAGGCTTTGTAGCTTTTCAAGCT 549
Qy 179 SerIleAspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGlu 198
Db 550 GCATTAATGCTGCTATCATGAAATCGCAAAATCATTCAGTATGATGAAACAGCTGATG 609
Qy 199 SerThrLysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArg 218
Db 610 TCAGTTACTGTTGATCATATGAATATTTACCTTTTGTGCCAAGGAGGTTCGAAC 669
Qy 219 GlyValIleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIle 238
Db 670 GATTTTTCATTTTCTTTTTCATTTTCTTTTCTATATATATATATATATATATATATAT 729
Qy 239 HisIleValAlaGluLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHis 258
Db 730 AATGTTACAAAGAAAGACAA---TACATTGCTCATGATGATGATGATGATGATGATGAT 786
Qy 259 AspThrAlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSer 278
Db 787 GAGTCAGCATTTCTGGCTTTCTGGGTTTGTATGATGCTGGCTTCATCTTATCATGGCC 846
Qy 279 LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleVal 298
Db 847 ACTTAAATGGCTCTTATTTGTAATAATGTCACAAATGTCGCTCTGCTGCTGCTGCTGATG 906
Qy 299 IlePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThr 318
Db 907 GTCTTCACCTCTTCTCTCTCTATGCTGCTCTTTGATAACTTTAGCTTTCTCTGATGAGT 966
Qy 319 ProLeuPheLysSerLysHisValGlyIleValGluPheValThrValAlaPhe 338
Db 967 GTGTTGATAAAGAAACCTTTCTCTTACGGGCTTGGTGTGTTCTCTCTCTCTCTCTTGG 1026
Qy 339 GlyPheIleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrp--- 357
Db 1027 GGGATCTGGGATTCCTCA---GCATTTGATATACACGCTCTCTGCTGCTTTTGAATGGACT 1083
Qy 358 ---LeuPheSerProPheCysHisCysThrPheValIleGlyIleAlaGlnValMet 375
Db 1084 TTGTGTCTTCTTAGCCCTTT-----GCCTTCACTGTGGGATGCCCGCAGCTTATA 1134
Qy 376 HisLeuGlu---AspPheAsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyr 394
Db 1135 CATTTGGACTATGATGTGATTTCTAATGCCCACTTGGATTCF---TCAAAATCCATAC 1191
Qy 395 ProLeuIleIleThrIleIleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAla 414
Db 1192 CTCATAATAGTACTCTTTTCTATGTTGTTTGTGACACCTTCTCTCTCTCTCTCTCTCT 1251
Qy 415 ValTyrLeuAspGlnValIleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPhe 434
Db 1252 TTATATTTTGAACAAATTTTGGCCGCTGAATATGACATCATGATGTTCTCCCTTGTTC 1311
Qy 435 LeuLysProSerTyrTrpSerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsn 454
Db 1312 CTGAAATCTGTTTGTGTTTCAACACAGGAGGCTAATCATGTGCTCTCTCTCTCTCTCT 1371
Qy 455 ValAsnGlyAsnIleSerPheSerGluIleIleGluProValSerSerGluPheValGly 474
Db 1372 ACAGATCTCTGATCTACATCAATGACTGTTTGTGAACAGTGTCTCCAGAAATTCCTGGG 1431

```

QY 848 AlaLysPheHisPheThrLeuLysArgGluSerLysSerValArgSerValLeuLeu 867  
 Db 3157 GCAACACTTCGCTTTCTTAAGATTAGGCGTGAAGAGAGCTCTTTTGTGTACTA 3216  
 QY 868 LeuLeuLeuLeuPhePheThrValGlnIlePheMetPheLeuValHisHisSerPheLys 887  
 Db 3217 GACTTGGAAATTGCTTTT---ATCCCATCTTCTAGAGAAAGATAATGTATAAAGTAACT 3273  
 QY 888 AenAlaValProIleLysLeuValProAspLeuTyPheLeuLysPheGlyAspLys 907  
 Db 3274 COTGAACTCAFTGTGGAGATTTTCACCCAGATATGTATTTCTTCTCTCGAACAATC 3333  
 QY 908 ProHisLysTyLysThrSerLeuLeuLeuGlnAenSerAlaAspSerAspIleSerAsp 927  
 Db 3334 CCGAAGACGCTCTTACCGAGCTGTTAATCGTTAATAATACAGGATCAAAATTATGAAGAC 3393  
 QY 928 LeuLeuSerPhePheThrSerGlnAenIleMetValThrMet-----Ile 942  
 Db 3394 CTGGTCATCTCACTGAAGTGTGAGATATAGTTTGGAAATAGATGACTTTAGAAACAGA 3453  
 QY 943 AsnAspSerAspTyLysValAlaProHisSerAlaAlaLeuAenValMetHisSer 962  
 Db 3454 AATGCTCAGATGATCCCTCC-----TACATGGAGCCATCATAGTGTCTGGTGAC 3504  
 QY 963 GluLysAspTyLysValPheAlaValPheAenSerThrMetValTySerLeuProIle 982  
 Db 3505 CAGAAGGATACAGATTTCTGTCGTGTAATACCAAGAAATGAATGTTTTCCTGTT 3564  
 QY 983 LeuValAsnIleIleSerAsnTyTyLeuTyHisLeuAenValThrGluThrIleGln 1002  
 Db 3565 CTATGGGAATTGTAGCATGCCCTATGGGAATTTTAACTTCACGGAGCTATTCAA 3624  
 QY 1003 IleTyPThrProPhePheGlnGluIleThrAspIleValPheLysIleLeuLeuTy 1022  
 Db 3625 ACGGAGAGCACTTATTTCTCGTAT-----GACATAGTCTGGATCTTGGTTTATA 3678  
 QY 1023 PheGlnAlaLeuLeuGlyIleValThrAlaMetProTyTyPheAlaMetGlu 1042  
 Db 3679 GATGGTCAATATTTTGTGTGTGATCACAACTGGTTCCTCTTTATCGCATGAGC 3738  
 QY 1043 AenAlaGluAenHisLysIleLysAlaTyThrGlnLeuLysLeuSerGlyLeuLeuPro 1062  
 Db 3739 AGCATCAGCGAATATAAAAAAATGTTCAATCCAGATTATGATTTTCAAGGCTCTGGCCT 3798  
 QY 1063 SerAlaTyTyPheGlyGlnAlaValAlaPheIleProLeuPhePheIleLeuLeu 1082  
 Db 3799 TCAGCATCTGTGTGGAGCAGCTCTGTGGACATTCATATATCTTCTGAT----- 3852  
 QY 1083 LeuMetLeuGlySerLeuLeuAlaPheHisTyGlyLeuTyPheTyThrValLysPhe 1102  
 Db 3853 -----CTCTTTCAATACATTAATTTACTATCTTCAATTTCTGGGATTC 3897  
 QY 1103 -----LeuAlaValValPheCysLeuLeuIleGlyTyValProSer 1115  
 Db 3898 CAGCTTTCAGGAACCTANGTTTGTGGTGTGATGATTAATGGTGTGACATTTCT 3957  
 QY 1116 ValIleLeuPheThrTyIleAlaSerPheThrPheLysLysIleLeuAenThrLysGlu 1135  
 Db 3958 CTATATCTCCTCAGATATGCTTTCATCTTTCGCAAGTGGCAAAAAAATAATGSC 4017  
 QY 1136 PheTyPThrPheLeuTySerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThr 1155  
 Db 4018 TTTTGGTCTTTTGGCTTTTAT---ATCTTAATATGATGTATCCACAAATATGATATCA 4074  
 QY 1156 PhePheMetGlyTyThrIleAlaThrIleLeuHisTyAlaPheCysIleIle---Ile 1174  
 Db 4075 -----ACTCAATATGAAAACTCACTTAATTTTGTGCATGATTTTCATA 4119  
 QY 1175 ProIleTyProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTyLysAenVal 1194  
 Db 4120 CCTTCCTTCACTTGTGGGTATGTCATGTTTATGATCCAGCTCGACTTT----- 4170

QY 1195 ArgLysAenValAspThrTyRAsnProTyRAspArgLeuSer----- 1208  
 Db 4171 ATGAGAACTTGGACAGTCTGGAC-----ANTAGAATAAATAGTCAATAAAACCAATT 4224  
 QY 1209 ---ValAlaValIleSerProTyRLeuGlnCysValLeuTyPhePheLeuGlnTyR 1227  
 Db 4225 CTTTAAACAACCTTAATACCATCTTACAGATGTTATTTCTTTTGTATAGGTGT 4284  
 QY 1228 TyRLeuLysTyRLeuGlyArgSerIleArgLysAspProPhePheArgAenLeuSer 1247  
 Db 4285 CTGGAATAGATGAAATGAATAATGAATAAAGACCCAGTTTTCAGA---ATCTCT 4341  
 QY 1248 ThrLysSerLysAenArgLysLeuProGluProProAspAenGluAspGluAspGluAsp 1267  
 Db 4342 CCAGGAGTAGAACAATCAT---CCCAATCGGAAGAGAGCCCGAAGAAAGATGAAGAT 4398  
 QY 1268 ValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGluLys 1287  
 Db 4399 GTTCAAGCTGAAAGAGTCCAGCAGCAATGCACTCACTGCTCCAAACTTGGAGGAGGAA 4458  
 QY 1288 ProSerIleMetValSerAenLeuHisLysGluTyRAspAspLysLysAspPheLeuLeu 1307  
 Db 4459 CCAGTCATCACTGCAAGCTGTTTACAAAGGAATATATGACACAAAGAAAGTTCCTTT 4518  
 QY 1308 SerArgLysValLysLysValAlaThrLysTyRLeuSerPheCysValLysLysGlyGlu 1327  
 Db 4519 TCACAGAAAGAAAGAAATAGCCATCAGAATGTTCTTTTGTGTTAAAGAGTGAA 4578  
 QY 1328 IleLeuGlyLeuLeuGlyProIleGlyAlaGlyLysSerThrIleIleAenIleLeuVal 1347  
 Db 4579 GTTTTGGGATTTACTAGACACAAATGAGCTGGTGTGTTAAAGTACTTCCATTAATAATACT 4638  
 QY 1348 GlyAspIleGluProThrSerGlyGlnValPheLeu---GlyAspTyRSerSerGluThr 1366  
 Db 4639 GGGTCACAAAGCCAACTCGAGAGTGGTGTGTACAGGCGACGACGACATCAGTAAGG 4698  
 QY 1367 SerGluAspAspSerLeuLysCysMetGlyTyRAspProGlnIleAenProLeuTyR 1386  
 Db 4699 CAACAGCATGACAAACGCTCAAGTCTTGGGGTACTGCGCTCAGGAGAACTCACTGTGG 4758  
 QY 1387 ProAspThrThrLeuGlnGluHisPheGluIleTyRLeuValLysGlyMetSerAla 1406  
 Db 4759 CCAGAGCTTACATGAAGAGCACTGGAGTGTGTATGACAGCTGTGAAGAGCTGGCAAA 4818  
 QY 1407 SerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeu 1426  
 Db 4819 GAAGATGCTGCTCTCAGTATTTTCAAGTGTGGAGCTCTTAAGCTCCAGGAACAATT 4878  
 QY 1427 GlnLysThrValLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSer 1446  
 Db 4879 AAGGCTCTGTGAAAACTCTATCAGAGGGAATAAAGAGAAAGCTGTGCTTGTGCTGAGC 4938  
 QY 1447 MetLeuGlyAenProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLys 1466  
 Db 4939 ATCTTGGGNAACCATCAGTGGTGTCTTAGATGAGCCGTTTCCCGGATGGACCCCGAG 4998  
 QY 1467 AlaLysGlnHisMetTyRArgAlaIleArgThrAlaPheLysAenArgLysArgAlaIle 1486  
 Db 4999 GGGCAGCAGCAAAATGTGGCAGATACCTGAGCTACCGTTTAAAAACAAGGAGGGGCACC 5058  
 QY 1487 IleLeuThrThrHisTyRMetGluGluAlaGluAlaValCysAspArgValAlaIleMet 1506  
 Db 5059 CTCTTGACCAACCATTCATATGTCAGAGCTGAGGCTGTGTGTGACCTGATGCCATGTG 5118  
 QY 1507 ValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLys 1526  
 Db 5119 GTGTGAGAAACGCTAAGGTGATTTGTTGCTCATCTCAACATCTCAAAACAAAGTTTGTAGA 5178  
 QY 1527 GlyTyRThrLeuGluIleLysLeuLysAspTyRTrpIleGluAenLeuGluValAspArgLeu 1546  
 Db 5179 GATTATTTTACTAGAAATAAAATGAAA-----GAACCTACCCAGGTGGAGCTCTC 5229  
 QY 1547 GlnArgGluIleGlnTyRThrIlePheProAenAlaSerArgGlnGluSerPheSerIle 1566





Db 1686 TCAACGAGTCTAAAAATAGAGAGCTTCCAGAACCCACACCAATGAGGATGAGATGAA 1745  
Qy 1267 AspVallylsAlaGluArgLeuLysVallylsGluLeuMetGlyCysGlnCysCysGluGlu 1286  
Db 1746 GATGCTAAAGCTGAAAGACTAAAGCTCAAGAGCTGATGGTGGCCAGTGTGAGGAG 1805  
Qy 1287 LysProSerileMetValSerAsnLeuHisGluGluTyrAspAspLysLysAspPheLeu 1306  
Db 1806 AAACCATCCATTATGGTCAGCAATTTGCATAAGATATGATGACAGAAAGATTTTCIT 1865  
Qy 1307 LeuSerArgLysVallylsValAlaThrLysTyrileSerPheCysVallylsGly 1326  
Db 1866 CTTTCAAGAAAGTAAAGAAAGTGGCAACTAAATACATCTCTTTCTGTGCAAAAGGA 1925  
Qy 1327 GluileLeuGlyLeuLeuGlyProAsnGluValadGlyLysSerThrileAsnileLeu 1346  
Db 1926 GAGATCTTAGACATATGGGTCGCAATGGTGCTGGCAAGACCAATTAATATCTG 1985  
Qy 1347 ValGlyAspilleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThr 1366  
Db 1986 GTTGGTGATATTGAACCAACTCAGGCCAGGTATTTTAGGAGATTAATCTTCAGAGACA 2045  
Qy 1367 SerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnileAsnProLeuTip 1386  
Db 2046 AGTGAAGATGATGTTCACTGAAGTGATGGTTACTGTCTCAGATPAACCCCTTGTGG 2105  
Qy 1387 ProAspThrThrLeuGlnGluHisPheGluileTyrGlyAlaVallylsGlyMetSerAla 1406  
Db 2106 CCAGATACATCTGACGAAACATTTTGAAATTTATGAGCTGTCAAGGAATGAGTGCA 2165  
Qy 1407 SerAspMetLysGluVallylsSerArgileThrHisAlaLeuAspLeuLysGluHisLeu 1426  
Db 2166 AGTGACATGAAGAGTCAATAGTCGAATAACACATGAGCTTGATTTAAAGAACATCTT 2225  
Qy 1427 GlnLysThrVallylsLeuProAlaGlyileLysArgLysLeuCysPheAlaLeuSer 1446  
Db 2226 CAGAAGACTGTAAGAAACTACCTGCAGGAATCAAAACGAAAGTTGTGTTGCTCTAAGT 2285  
Qy 1447 MetLeuGlyAsnProGlnileThrLeuLeuAspGluProSerThrGlyMetAspProLys 1466  
Db 2286 ATGCTAGGGAATCTCTCAGATTAATTTGTAGTAGAACCATCTACAGGTATGGATCCCAAA 2345  
Qy 1467 AlaLysGlnHisMetTyrArgAlaileArgThrAlaPheLysAsnArgLysArgAlaAla 1486  
Db 2346 GCCAAACAGCACATGTCGCGAGCATTCGAATGCAATTTAAACAGAAACGCGGTGCT 2405  
Qy 1487 IleLeuThrThrHisTyrMetGluAlaGluAlaValCysAspArgValAlaileMet 1506  
Db 2406 ATTGACCACTCACTATATGAGAGAGCAGAGGCTGTCTGTGATCGAGTAGCTATCATG 2465  
Qy 1507 ValSerGlyGlnLeuArgCysileGlyThrValGlnHisLeuLysSerLysPheGlyLys 1526  
Db 2466 GTGCTGGGCACTTAAGATGATATCGAACAGCTACAAACATCTAAGAGTAGATTTGGAAA 2525  
Qy 1527 GlyTyrPheLeuGluileLysLeuLysAspTyrilleGluAsnLeuGluValAspArgLeu 1546  
Db 2526 GGCTACTTTTGGAAATTAATTTGAAGGACTGATAGAAAACCTAGAAAGTAGACCGCTT 2585  
Qy 1547 GlnArgGluileGlnTyrilePheProAsnAlaSerArgGlnGluSerPheSerSerile 1566  
Db 2586 CAAAGAGAAATTCAGTATATTTCCCAATGACCCGTGAGGAAGTTTCTCTATT 2645  
Qy 1567 LeuAlaTyrLysileProLysGluAspValGlnSerLeuSerGlnSerPheLysLeu 1586  
Db 2646 TTGGCTTATAAATTCCTAAGCAAGATGTTCAAGTCCCTTCAACATCTTTTAAAGCTG 2705  
Qy 1587 GluGluAlaLysHisAlaPheAlaileGluGluTyrSerPheSerGlnAlaThrleuGlu 1606  
Db 2706 GAAGAGCTTAACATGCTTTTCCCAATGAAAGATATAGCTTTCTCAGGCAACATTTGAA 2765  
Qy 1607 GlnValPheValGluLeuThrLysGluGlnGluGluAspAsnSerCysGlyThrLeu 1626  
Db 2766 CAGGTTTTGTAGACTCACTAAAGACACAGAGAGGAGGAATATAGTTGTGGAACCTTTA 2825

Qy 1627 AnSerThrLeuTyrTrpGluArgThrGlnGluAspArgValValPhe 1642  
Db 2826 AACACACACTTTGGTGGAAACGACACACAGAGATAGTAGTATTT 2873

RESULT 10

US-10-005-338B-4  
Sequence 4, Application US/10005338B  
Publication No. US20030044895A1  
GENERAL INFORMATION:  
APPLICANT: DENEELE, Patrice  
APPLICANT: ROSIER-MONTUS, Marie-Francoise  
APPLICANT: PRADRES, Catherine  
APPLICANT: ARNOULD-REGUIGNE, Isabelle  
APPLICANT: DUVERGER, Nicolas  
APPLICANT: ALLIKMETS, Rando  
APPLICANT: DEAN, Michael  
TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENE  
TITLE OF INVENTION: CONTAINING SUCH NUCLEIC ACIDS, AND USES THEREOF  
FILE REFERENCE: ABCA5, 6, 9, 10  
CURRENT APPLICATION NUMBER: US/10/005,338B  
CURRENT FILING DATE: 2001-12-07  
PRIOR APPLICATION NUMBER: US 60/263,231  
PRIOR FILING DATE: 2001-01-23  
PRIOR APPLICATION NUMBER: FR 00403440.1  
PRIOR FILING DATE: 2000-12-07  
NUMBER OF SEQ ID NOS: 217  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 6181  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 1420  
OTHER INFORMATION: n=unknown, may be a or g or c or t  
US-10-005-338B-4

Alignment Scores:  
Pred. No.: 1,11e-254 Length: 6181  
Score: 3234.00 Matches: 714  
Percent Similarity: 61.39% Conservative: 313  
Best Local Similarity: 42.68% Mismatches: 561  
Query Match: 38.38% Indels: 86  
DB: 14 Gaps: 30

US-10-090-458-5 (1-1642) x US-10-005-338B-4 (1-6181)

Qy 6 ArgGluValGlyValTyrpArgGlnThrArgThrLeuLeuLysAsnTyrileuLys 25  
Db 624 AAAGAGATAAGCGTGGCTCAACAAATTCAGGCTCTCTGTACAGAAATTTCTTAAAAA 683  
Qy 26 CysArgThrLysLysSerSerValGlnGluLeuLeuPheProLeuPhePheLeuPheTrp 45  
Db 684 TGGAGATATAAAGAG--AGTTATTGGATGACACATAACATTGTTTCTAGGGCTATAT 741  
Qy 46 LeuileLeuileSerMetMetHisProAsnLysLysTyrGluGluValProAsnileGlu 65  
Db 742 TTGTGATCTTTTCGGAACACTTCAGACTACCCGTTTTCCTGAACACCTCTCTAAAGTC 801  
Qy 66 LeuAsnProMetAspLysPheThrLeuSerAsnLeuileLeuGlyTyrThrProValThr 85  
Db 802 CTGGGAAGCGTGGATCAGTTTAACTGCTGCGCTGGTAGTGGCATATACACCACTCACT 861  
Qy 86 AsnileThrSerSerileMetGlnLysValSerThrAspHisLeu-----ProAsp 102  
Db 862 AACATAACACAAAGGATAATGAATAAGATGGCTTGGCTTTCCTTTATGAAGAGGAGACA 921  
Qy 103 ValileileThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLys 122  
Db 922 GTCATTTGGGACACCAT-----GAAGAGACCATGGATATAGAACTTCCAAAAA 972  
Qy 123 ProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArgPhe--- 141

APPLICANT: TAMECHIKA, ICHIRO  
 APPLICANT: SEKI, NAOHICO  
 APPLICANT: YOSHIKAWA, TSUTOMU  
 APPLICANT: OTSUKA, MOTOKU  
 APPLICANT: NAGAHARI, KENJI  
 APPLICANT: MASUHO, YASUHIKO  
 TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA  
 FILE REFERENCE: 084335/0160  
 CURRENT APPLICATION NUMBER: US/10/094, 749  
 CURRENT FILING DATE: 2002-03-12  
 PRIOR APPLICATION NUMBER: 60/350,435  
 PRIOR FILING DATE: 2002-01-24  
 PRIOR APPLICATION NUMBER: JP 2001-328381  
 PRIOR FILING DATE: 2001-09-14  
 NUMBER OF SEQ ID NOS: 3381  
 SOFTWARE: PatentIn ver. 2.1  
 SEQ ID NO 984  
 LENGTH: 3268  
 TYPE: DNA  
 ORGANISM: Homo sapiens  
 US-10-094-749-984

Alignment Scores:  
 Pred. No.: 0 Length: 3268  
 Score: 4847.50 Matches: 941  
 Percent Similarity: 98.64% Conservative: 2  
 Best Local Similarity: 98.43% Mismatches: 2  
 Query Match: 57.53% Indels: 11  
 DB: 15 Gaps: 1

US-10-090-458-5 (1-1642) x US-10-094-749-984 (1-3268)

698 MetPheLeuLysSerLysTyrGlyLeuValArgLeuSerMetTyrIleAspLysTyr 717  
 6 ATGTCCTCAAAAGTAAATGGGATCGGTACCGCTGAGCATGTACATAGACAAATAT 65  
 718 CysAlaThrGluSerLeuSerLeuValLysGlnHisIleProGlyAlaThrLeuLeu 737  
 66 TGTGCCACAGAAATCTCTTTCTTCTACTGTTTAAACAAACATATACCTGGAGCTACTTTATTA 125  
 738 GlnGlnAsnAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSer 757  
 126 CAACAGATGACCAACAACTTGTGTATAGCTTGCCTTCAAGGACATGGACAAATTTTCA 185  
 758 Gly-----LeuPheSerAlaLeuAspSerHis 766  
 186 GGAATTCCTTGATAGACAAAAGAGATGTTATGTATGTTTCTGCTGAGACAGTCAT 245  
 767 SerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeu 786  
 246 TCATAATTTGGGTGTCATTTCTTATGGTGTTCATGACGACTTTGGAGACGTATTTTA 305  
 787 LysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnPro 806  
 306 AAGCTAGAAAGTTGAAGCAGAAATGACCAAGCAGATATAGTGTATTTACTCAGCAGCCA 365  
 807 LeuGluGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuLeu 826  
 366 CTGGAGGAAGAAATGGATTCAAATCTTTTGTGAAATGGAAAGACAGCTTACTTATCTT 425  
 827 SerGluThrLysAlaSerLeuValSerThrMetSerLeuThrLysGlnGlnMetTyrThr 846  
 426 TCTGAACCAAGGCTTCTTAGTGAGCACCATGAGCCTTTGAAACCAACAGATGTATACA 485  
 847 IleAlaLysPheHisPheThrLeuLysArgGluSerLysSerValArgSerValLeu 866  
 486 ATAGCAAAAGTTTCATTTCTTACCTTGAAACGTGAAGATTAATCAGTGAGATCAGTGTG 545  
 867 LeuLeuLeuLeuLeuPhePheThrValGlnIlePheMetPheLeuValHisHisSerPhe 886  
 546 CTTCCTGCTTTAAATTTTTCACAGTTTCAGATTTTATGTTTTTGGTTTCATCTCTTT 605  
 887 LysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLysPheGlyAsp 906

606 AAAATGCTGTGGTCCCATCAAACTGTTCCAGACTTATATTTCTAAACCTGGAGAC 665  
 907 LysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAspIleSer 926  
 666 AAACCAACATAAATACAAACCAAGTCTGCTTCTTCAAAATCTGCTGACTCAGATATCAGT 725  
 927 AspLeuIleSerPhePheThrSerGlnAsnIleMetValThrMetIleAsnAspSerAsp 946  
 726 GATCTTATAGCTTTTTCACAGCCAGACACATAATGTTGTCAGATGATTAATGACAGTAC 785  
 947 TyrValSerValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyr 966  
 786 TATGTATCCGTGGCTCCCATATGATGCTGCTTAAATGTGATGCATTCAGAAAGACATAT 845  
 967 ValPheAlaAlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIle 986  
 846 GTTTTGGAGCTGTTTTCAACAGTACTATGTTTATCTTTTACCTATATATAGTATATC 905  
 987 IleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTyrSerThr 1006  
 906 ATTAGTAACACTACTATCTTTATCAATTAATGTGACTGAACCATCCAGATCTGGAGTACC 965  
 1007 ProPhePheGlnGluIleThrAspIleValPheLysIleGluLeuTyrPheGlnAlaAla 1026  
 966 CCATTCCTTTCAAGAAATTTACTGATATAGTTTTTAAATGAGCTGTATTTTCAACAGCT 1025  
 1027 LeuLeuGlyIleIleValThrAlaMetProTyrPheAlaMetGluAsnAlaGluAsn 1046  
 1026 TTGCTTGGAAATCATTTGTTACTGCAATGCCACTTACTTTGCCATGGAAATGCGAGAAAT 1085  
 1047 HisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrp 1066  
 1086 CATNAGATCAAAAGCTTTTACTCAACTTAACTTTTCAGGTCTTTTGCCATGTCATATGG 1145  
 1067 IleGlyGlnAlaValAlaAspIleProLeuPhePheIleIleLeuIleLeuMetLeuGly 1086  
 1146 ATTGACACAGCTGTTGTTGATATCCCTTATTTTATCATCTTATTTGATGTAGGA 1205  
 1087 SerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaValVal 1106  
 1206 AGCTTACTGGCATTTTCAATATGGATTTATTTTATATCTGTAAAGTTCCTTGTGCTGGT 1265  
 1107 PheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPheThr 1126  
 1266 TTTTGGCTTATGGTATGTTTCCATCAGTATTCGTTCCTTATATATGCTTCTTTCACC 1325  
 1127 PheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyrSerValAlaAlaLeu 1146  
 1326 TTTAAGAAATTTTAAATACCAAGAAATTTGGTCAATTTATCTATCTTCTGTGACGCTG 1385  
 1147 AlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyrThrIleAlaThrIleLeu 1166  
 1386 GCTGTATTTGCAATCAGTGAATAAATCTTTTATGGATACACAAATTTGCACTATTTCTT 1445  
 1167 HisTyrAlaPheCysIleIleIleProLeuTyrProLeuLeuGlyCysLeuIleSerPhe 1186  
 1446 CATATGCTTTTGTATCATCATCTTCAATCTATCCACTTCTAGGTTGCTGATTTCTTTC 1505  
 1187 IleLysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnProTrpAspArg 1206  
 1506 ATAAAGATTTCTTGAAGAAATGTACGAAATAATGGACACCTATATAATCCATGGGATAGG 1565  
 1207 LeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGln 1226  
 1566 CTTTCAGTACTGTTTATATCGCTCACCTGAGTGTGACTGTGATGCTGATTTCTCTTCAAA 1625  
 1227 TyrTyrGluLysLysTyrGlyArgSerIleArgLysAspProphePheArgAsnLeu 1246  
 1626 TACTATGAGAAAAATATGAGCGAGATCAATGAAGAAAGATCCCTTTTTCAGAAACCTT 1685  
 1247 SerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAspGluAspGlu 1266

QY	1018	siileGluLeuTyrrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTy	1038	QY	1378	rCyaproglnileAsnProLeuTrpProAspThrThrLeuGlnGlnHisPheGluLeuTy	1398
Db	1082	AATTGAGCTGTAATTTTCAGCAGCTTTCCTGGAATCATTTGTTACTGCAATGCACCTTA	1141	Db	2131	CTGTCCTCAGATAAACCCTTTGTGCGCAGATACATACATTCGAGGAACATTTTGAATTTA	2190
QY	1038	rPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrrThrGlnLeuLysLeuSe	1058	QY	1398	rGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHi	1418
Db	1142	CTTTGCCATGAAATGCAGAAATCATAA-----	1171	Db	2191	TGGAGCTGTCAAGGAATGAGTGCAGATCAATGAAAGAGTCATAAGTGGAAATACACA	2250
QY	1058	rGlyLeuLeuProSerAlaTyrrPheGlyGlnAlaValAlaAspIleProLeuPhePh	1078	QY	1418	salAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLy	1438
Db	1172	-GGTCTTTGGCATCTGCATATTGATGGACAGCTGTTGTTGATATCCCTTATTTT	1230	Db	2251	TGCATCTGATTTTAAAGAACATCTTCAGAGACTGTAAGAAACTACCTCGCAGGAATCAA	2310
QY	1078	eIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisIleGlyLeuTyrrPheTy	1098	QY	1438	sArgLysLeuCyusPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGI	1458
Db	1231	TATCATCTTATTTGATGCTAGAGCTTACTGGCATTTTCATATGATGATATATTTTA	1290	Db	2311	ACGAAAGTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGA	2370
QY	1098	rThrValLysPheLeuAlaValAlaPheCysLeuIleGlyTyrrValProSerValIleLe	1118	QY	1458	uProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAl	1478
Db	1291	TACTGTAAGTTCCTTGCTGTGGTTTTTCCTTATTTGGTTATGTTTATGTCATCAGTTATCT	1350	Db	2371	ACCATCTCAGATATGATATGATCCCAAGCCAAACAGACACATGTGGCGAGCAATTCGAACGC	2430
QY	1118	uPheThrTyrrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSe	1138	QY	1478	aPheLysAsnArgLysArgAlaAlaIleLeuThrThrHisTyrrMetGluGluAlaGluAl	1498
Db	1351	GTTCACTATATTTGCTTTCACCTTTAAGAAAAATTTTAAATACCAAGAAATTTTGGTC	1410	Db	2431	ATTTAAAAACAGAAAGCGGCTGTATTTCTGACCACTCCTATATATGAGGAGGAGAGGC	2490
QY	1138	rPheIleTyrrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMe	1158	QY	1498	aValCysAspArgValAlaAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGI	1518
Db	1411	ATTATATCATTTCTGGCAGCGCTGGCTGTGATTTGCAATCATCTGAATAAATCTTTTAT	1470	Db	2491	TGTCGTGATCAGATAGCTATCATGTGCTGGCAGCTTAAGATGTATCGGACACAGTACA	2550
QY	1158	tGlyTyrrThrIleAlaThrIleLeuHisTyrrAlaPheCysIleIleIleProIleTyrrPr	1178	QY	1518	nHisLeuLysSerLysPheGlyLysGlyTyrrPheLeuGluIleLysLeuLysAspTrpII	1538
Db	1471	GGGATACACAATTCGAATCTTCTTCAATATGCTTTCCTTTTGTATCATCATTCGAATCTATCC	1530	Db	2551	ACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTTAAATTTGAAGGACTGGAT	2610
QY	1178	oLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnVa	1198	QY	1538	eGluAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrrIlePheProAsnAlaSe	1558
Db	1531	ACTCTAGTGTCCGTAATTTCTTCAATAAATTTCTTGGAGAGATGTACGAAAAATGT	1590	Db	2611	AGAAAACTTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATTCGAAG	2670
QY	1198	lAspThrTyrrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrrLeuGlnCy	1218	QY	1558	rArgGlnGluSerPheSerSerIleLeuAlaTyrrLysIleProLysGluAspValGlnSe	1578
Db	1591	GGACACCTATAATCCATGGATAGGCTTTCAGTAGCTGTATATCGCTTACCTGCGAGTG	1650	Db	2671	CCGTCAGGAAAGTTCCTTCTTATTTTGGCTCATAAATTCCTAAGGAAGATGTTTCAGTC	2730
QY	1218	sValLeuTrpIlePheLeuLeuGlnTyrrTyrrGluLysIleTyrrGlyArgSerIleAr	1238	QY	1578	rLeuSerGlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTy	1598
Db	1651	TGTACTGTGGATTTTCTCTTCAATACATCTATGAGAAAAAATATGAGGACGATCANTAG	1710	Db	2731	CTTTTCACAATCTTTTTTAAAGCTGGAAGAACTAAACATGCTTTTGCCTTTGAAGAATA	2790
QY	1238	gLyAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluPr	1258	QY	1598	rSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGI	1618
Db	1711	AAAAGATCCCTTTTCAGAAACCTTTCAAGGAAGTCTAAAAATAGGAAGCTTCCAGAACC	1770	Db	2791	TAGCTTTTCTCAAGCAACATTCGAAACAGGTTTTTGTAGAACTCCTTAAAGAACAGAGA	2850
QY	1258	oProAspAsnGluAspGluAspValLysAlaGluArgLeuLysValLysGluLe	1278	QY	1618	uGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAs	1638
Db	1771	ACCAGACAATGAGGATGAAGATGAAGATGCAAGCTGAAGACTAAAGGTCAAGAGCT	1830	Db	2851	GGAGATAATAGTTGTGGAACCTTTAAACAGACACACTTTGTGGGAACGAAACCAAGAAGA	2910
QY	1278	uMetGlyCysGlnCysGluGluLysProSerIleMetValSerAsnLeuHisLysGI	1298	QY	1638	parGValValphe	1642
Db	1831	GATGGGTGCCAGTGTGTGAGAGAGAAACCATCCATTATGCTCAGCAATTTGCATAAAGA	1890	Db	2911	TAGAGTAGTATTT	2923
QY	1298	uTyrrAspAspLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTy	1318	RESULT 9			
Db	1891	ATATGATGACAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTGGCAACTAAATA	1950	US-10-094-749-984			
QY	1318	rIleSerPheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGI	1338	; Sequence 984, Application US/10094749			
Db	1951	CATCTCTTTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGGTCCAAATGGTCTGG	2010	; Publication No. US20030219741A1			
QY	1338	yLysSerThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPh	1358	; GENERAL INFORMATION:			
Db	2011	CAAAAGCACAAATTAATATCTGTTGTTGTTGATATGAAACCAACTTCAGGCGCAGGTATT	2070	; APPLICANT: ISOGAI, TAKAO			
QY	1358	eLeuGlyAspTyrrSerGluThrSerGluAspAspSerLeuLysCysMetGlyTy	1378	; APPLICANT: SUGIYAMA, TOMOYASU			
Db	2071	TTTAGGAGATTTTCTTCAGAGACAAGTGAAGATGATGATTCATCTGAGTGTATGGTTA	2130	; APPLICANT: OTSUKI, TETSUJI			
				; APPLICANT: WAKAMATSU, AI			
				; APPLICANT: SATO, HIROYUKI			
				; APPLICANT: ISHII, SHIZUKO			
				; APPLICANT: YAMAMOTO, JUN-ICHI			
				; APPLICANT: ISONO, YUUKO			
				; APPLICANT: HIO, YURI			
				; APPLICANT: OTSUKA, KAORU			
				; APPLICANT: NAGAI, KEIICHI			
				; APPLICANT: IRIE, RYOTARO			

4081 GATTATCTTCAGAGCAAGTGAAGATGATGATTCCTGAGTGTATGGCTTACTGTCT 4140  
 1381 GlnIleAsnProLeuTTPProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
 4141 CAGATAAACCCCTTTGGCCGAGATACATACATTCGAGAAACATTTTGAAATTTATGAGCT 4200  
 1401 ValIysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
 4201 GTCAAGGATGAGTGCAGAGTACATGAAGAAGTCATAAGTCGAATTAACACATGCACCT 4260  
 1421 AspLeuLysGluHisLeuGlnIleThrValIysLysLeuProAlaGlyIleLysArgLys 1440  
 4261 GATTATAAGAACATCTTCAGAACCTGTAAGAAGACTACCTGCAGGAATCAACAGGAAG 4320  
 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
 4321 TTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACCATCT 4380  
 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
 4381 ACAGTATGATGCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAACCTGCATTTAAA 4440  
 1481 AsnArgLysArgAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCys 1500  
 4441 AACAGAAAGCCGGCTGCTATTCTGACCACTCAGTATATGGAGAGGCGAGGCTGTCTGT 4500  
 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
 4501 GATCGAGTAGCTATCATGTGCTGGCGAGTTAAGATGTATCGGAACAGTACAACTCTA 4560  
 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
 4561 AAGAGTAAATTTGAAAGAGGCTACTTTTGGAAATTTAAATGAAGGACTGGATAGAAAC 4620  
 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560  
 4621 CTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGAAGCCGTCAG 4680  
 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580  
 4681 GAAAGTTTCTCTTCTATTGCTTATATAAATTCCTAAGGAAGATGTTTCAGTCCCTTTCA 4740  
 1581 GlnSerPhePheLeuGluGlu 1588  
 4741 CAATCTTTTAACTGCGAGAA 4764

RESULT 8  
 US-10-108-260A-160  
 ; Sequence 160, Application US/10108260A  
 ; Publication No. US20040005560A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: HELIX RESEARCH INSTITUTE  
 ; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA  
 ; FILE REFERENCE: H1-A0106  
 ; CURRENT APPLICATION NUMBER: US/10/108,260A  
 ; CURRENT FILING DATE: 2002-03-27  
 ; NUMBER OF SEQ ID NOS: 5458  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 160  
 ; LENGTH: 3347  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-10-108-260A-160

Alignment Scores:  
 Pred. No.: 0 Length: 3347  
 Score: 4930.50 Matches: 966  
 Percent Similarity: 98.17% Conservative: 1  
 Best Local Similarity: 98.07% Mismatches: 2  
 Query Match: 58.52% Indels: 17  
 DB: 15 Gaps: 2

US-10-090-458-5 (1-1642) x US-10-108-260A-160 (1-3347)  
 664 ArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLys 683  
 2 CGGGTGACAGTGTTCAGTACTCAITTCATGGATGAAGCTGACATTCCTTTCAGATAGGAAA 61  
 684 AlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLys 703  
 62 GCTGTGATATCAAGGAATGCTGAATGTGTGGTTCTTCAATGTTCTCTCAAAAGTAAA 121  
 704 TrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeu 723  
 122 TGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACAGATCTCTT 181  
 724 SerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGln 743  
 182 TCTTCACCTGGTTAAACAACATATACCTGGAGCTACTTTATTACACAGAAATGACCAAA 241  
 744 LeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeu 763  
 242 CTGTGTATAGCTTGCTTTCAAGCATGGACAAATTTTCAGGTTTGTCTTCTGCCCTA 301  
 764 AspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAsp 783  
 302 GACAGTCATTCAAATTTGGGTGTCTTCTTATGGTGTCTTCCATGACGACTTTGGAAGAC 361  
 784 ValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThr 803  
 362 GTATTTTAAAGCTAGAAGTTGAAGCAGAAATTTGACCAAGCAGATATAGTGTATTTACT 421  
 804 GlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu 823  
 422 CAGCAGCCACTGGAGGAGAAATGGATTCAAATCTTTTGTATGAATGAACAGAGCTTA 481  
 824 LeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGln 843  
 482 CTATTTCTTTTGAACCAAGCTTCTCTAGTGAAGCAGCATGAGGCTTTGGAAACAAACAG 541  
 844 MetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLysSerValArg 863  
 542 ARGATATCAATAGCAAGTTTCATTTCTTACCTTGAACCGTGAAGTAAATCAGTGAGA 601  
 864 SerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPheLeuValHis 883  
 602 TCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTTGTGTTAT 661  
 884 HisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyrPheLeuLys 903  
 662 CACTCTTTTAAAAATGCTGTGGTTCCTCAAACTTGTTCAGACTTATATTTTCTAAA 721  
 904 ProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAla----- 921  
 722 CCTGGAGACAAACACATATAACAAACAGTCTGCTTCTTCAAAATTCCTGCTGTGAG 781  
 922 -----AspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleW 938  
 782 AGTGTGTGAAGACTCAGATATCTGTGATCTTATTAGCTTTTTCACAGCAGCAACATAT 841  
 938 tValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAs 958  
 842 GGTGACGATGATTAATGACAGTGACTATGTATCCGTGGCTCCCATAGTGGCGCTTAAA 901  
 958 nValMetHisSerGluLysAspTyrValPheAlaValPheAsnSerThrMetValTyr 978  
 902 TGTGATGCTATCAGAAAGGACTATGTTTTTGCAGCTGTTTTTCAACAGTACTATGTTTAA 961  
 978 rSerLeuProIleLeuValAsnIleSerAsnTyrTyrLeuTyrHisLeuAsnValTrn 998  
 962 TTCTTTTACCTATATAGTGAATATCATTAGTAGTAACTACTATCTTTTATCAATTAATGTGAC 1021  
 998 rGluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLys 1018  
 1022 TGAACACCTCCAGATCTGGAGTACCCCATTTCTTCAAGAAATTTACTGATATAGTTTTTAA 1081

QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
DB 1921 CCAACAGCTGGAAATGGACCCCTGTTCTCGACATATTGATGGAATCTTTTAAAAATCAGA 1980  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
DB 1981 AAAGCCAATCGGTGACAGTGTTCAGTACTCATTTCAATTCATGGAAGCTGACATCTCTGCA 2040  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
DB 2041 CATAGGAAGCTGTGATATCACAGGAATGCTGAATGTGTGTTCTTCAATGTTCTC 2100  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 2101 AAAAGTAAATGGGGATCGGTACCGCTGACATGATGACATAGACAAATATTGTCACCA 2160  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsn 740  
DB 2161 GAATCTCTTTCTCTCTGTTAAACAACATATATCTGGAGTACTTTTATACACAGAAT 2220  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
DB 2221 GACCAACAACATGTGTATAGTTCCTTTCAAGGACATGACAAATTTTCAGGTTCCTTT 2280  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 2281 TCTGCCCTAGACAGTCATTCAAATTTGGTGGCATTTCTTAGGGTTTCCATGACCACT 2340  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
DB 2341 TTGGAAGACGTATTTTAAAGCTAGAAGTTCAAGCAGAAATGACCAAGCAGATATAGT 2400  
QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 2401 GTATTACTACGACGCCCTGGAGGAATGGAATGCAAAATCTTTTGATGAATGGA 2460  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
DB 2461 CAGAGCTTACTTATTTCTGAACCAAGGCTKCTCTAGTGAGCACCATGAGCCTTTGG 2520  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
DB 2521 AAACAACAGATGATATACAAATAGCAAGTTTCATTTCTTTACCTTCAACACGFGAAAGTAA 2580  
QY 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 2581 TCAGTGAGATCAGTGTGCTTCTGCTTTTAATTTTTCAGATTTCAGATTTTATGTTT 2640  
QY 881 LeuValHisHisSerPheLysAsnAlaValProIleLysLeuValProAspLeuTyr 900  
DB 2641 TTGGTTCATCACTCTTTTAAAAATGCTGTGGTTCCCATCAAACTGTTTCCAGACTTATAT 2700  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
DB 2701 TTTCTAAAACCTGGAGACCAACACATAAATACAAAACAAGTCGTCTTCAAAATCT 2760  
QY 921 AlaAspSerAspIleSerAspLeuLeuSerPhePheThrSerGlnAsnIleMetValThr 940  
DB 2761 GCTGACTCAGATATCAGTATCTTATTAGCTTTTTCACAAGCCAGAACATATGCTGACG 2820  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
DB 2821 ATGATTATGACAGTGACATGATGATCCGTGGCTCCCATAGTGGCTTTTAAATGTGTG 2880  
QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
DB 2881 CATTCAGAAAAGCACTATGTTTTCAGCTGTGTTTCAACAGTACTATGTTTATCTTTA 2940  
QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
DB 2941 CCTATATTAGTGAATATCATTAGTAACACTATCTTTATCATTTTAAATGTGACTGAAC 3000

QY 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
DB 3001 ATCCAGATCTGGAGTACCCCATCTTTCAAGAAATTTACTGATATAGTTTTTAAAAATTGAG 3060  
QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
DB 3061 CTGTATTTTCAAGCAGCTTTCCTTGGAAATCATTTGTTACTGCAATGCCACCTTACTTTGCC 3120  
QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
DB 3121 ATGGAAAATGACAGAAATCATAAGATCAAGCTTAYACTCAACTTAAACTTTTCAGGCTCT 3180  
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLeuPhePheIleIle 1080  
DB 3181 TTGCCATCTGCATATTTGGATTGGACAGCTGTTGTTGATATCCCTTATTTTATCATCT 3240  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1100  
DB 3241 CTATTTTTCATCTAGGAAGCTTATTGGCATTTTCATTATGATGATATATTTTATACTGTA 3300  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
DB 3301 AAGTTCCTTCTGCTGCTGTTTTTGGCTTATTTGGTTATGTTCCATCAGTTATTTCTGTTCACT 3360  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
DB 3361 TATATTGCTTCTTCCACCTTTAAGAAAAATTTTAAATACCAAGAAATTTTGGTCACTTTATC 3420  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
DB 3421 TATTCCTGGCAGCGTGGCTGTATTGCAATCATCTGAATAAATCTTTCTTTATGGATAC 3480  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
DB 3481 ACAATTGCAACTATTCTTTCATTTATGCTTTGATCATCATCTTCCAATCTATCCACTTCTA 3540  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
DB 3541 GGTGCTGCTGATTCTTTTCATAAAGATTTCTTGAAGAATGTACGAAAAAATGTGACACC 3600  
QY 1201 TyrAsnProTyrAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
DB 3601 TATATATCATGGATAGGCTTTTCAGTAGCTGTATATTCGCTTACCTGAGTGTGTA 3660  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
DB 3661 TGGATTTTCTCTTACATATCTATGAGAAAAATATGAGGCGAGATCAATAAGAAAAAGAT 3720  
QY 1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAsp 1260  
DB 3721 CCTTTTTTCAGAAACCTTTTCAACGAAGCTTAAAAATAGGAAGCTTCCAGAACCCACAGAC 3780  
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
DB 3781 AATGAGATGAAGATGAGATGCTCAAAAGCTGAAAGCTAAAGGTCAAGAGCTGATGGGT 3840  
QY 1281 CysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
DB 3841 TGGCAGTGTGTGAGGAGAAACCATCTTATTTGGTCAGCAATTTGCATTAAGAAATATGAT 3900  
QY 1301 AspLysLeuAspPheLeuLeuSerArgLysValLysValLysValAlaThrLysTyrIleSer 1320  
DB 3901 GACAAGAAAGATTTTCTTCTTCAAGAAAAAGTAAAGAAAGTGGCAACTTAAATACATCTCT 3960  
QY 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValAlaGlyLysSer 1340  
DB 3961 TTTCTGTGAAAAAAGAGAGATCTTAGGACTATTGGGTCCCAATTTGGTCTGGCAAAAGC 4020  
QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
DB 4021 ACAATTATTAAATTTCTGTTGTTGATATTGAACCAACTTTCAGGCCAGGTATTTTATGGA 4080  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380

;; PRIOR FILING DATE: 2000-10-10  
;; NUMBER OF SEQ ID NOS: 5  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 3  
;; LENGTH: 4785  
;; TYPE: DNA  
;; ORGANISM: homo sapiens  
US-09-971-121-3

Alignment Scores:  
Pred. No.: 0 Length: 4785  
Score: 8112.00 Matches: 1583  
Percent Similarity: 99.69% Conservative: 0  
Best Local Similarity: 99.69% Mismatches: 5  
Query Match: 96.27% Indels: 0  
DB: 9 Gaps: 0

US-10-090-458-5 (1-1642) x US-09-971-121-3 (1-4785)

```
Qy 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20
Db 1 ATGCCACTGCAATTAGGGAGTAGGAGTTTGGAGACAGACAGACAGACACTTCTACTGAAG 60

Qy 21 AsnTyrLeuIleLysCysArgThrIleLysSerValGlnGluLeuLeuPheProLeu 40
Db 61 AATTACTTAAATTAATGAGAACCAAAAGAGTAGTGTTCAGGAATTCITTTTCCACTA 120

Qy 41 PhePheLeuPheTrpLeuIleLeuSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 121 TTTTATTTTATTTGGTTAATTAATTAATAGCATGATGCATCCAAATAAGAAATAGAA 180

Qy 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80
Db 181 GTGCCATAATAGAACTCAATCTTAAGGACAGATTACTCTTCTTAATTAATTTCTTGG 240

Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
Db 241 TATACTCCAGTGACATAATATTACAGCAGCATCATCGAGAAAGTGTCTACTGATCACTA 300

Qy 101 ProAspValIleIleLeuThrGluGluTyrThrAsnGlnLysGluMetLeuThrSerSerLeu 120
Db 301 CCTGATGTCATAATTAATGAGAAATATACAAATGAAAGAAATGTTAAACATCCAGTCTC 360

Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
Db 361 TCTAGCCGAGCAACTTTGTAGGTGTGTTTCAAGAGCTCCATGCTCTATGAACTTCGT 420

Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
Db 421 TTTTTCCTGATATGATCCAGTATCTTCTATTATATATGATGATTCAGAGCTGGGTGTTCA 480

Qy 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180
Db 481 AAATCATGTGAGGCTGCTCAGTACTGGTCTCCTCAGGTTTCAGGTTTACAGCATCCATA 540

Qy 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200
Db 541 GATGCTGCCATATATACAGTTCAAGACCAATGTTCTCTTTGGAAGGAGCTGGAGTCAACT 600

Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
Db 601 AAAGCTGTTATATGGGAAAGAACTGCTGTTGTAGAAATAGATATACCTTTTCCCGAGGAGTA 660

Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
Db 661 ATTTTAAATATACCTAGTTATAGCATTTTCACCTTTTGGATACCTTTTGGCAATTCATATC 720

Qy 241 ValAlaGlnLysGluLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
Db 721 GTAGCAGAAAGAAAGAAATATAAGAAATTTTAAAGATAATGGGACTTCATGATACT 780

Qy 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280
```

```
Db 781 GCCTTTTGGCTTCTCTGGGTTCTTCTATAWACAAGTTTAAATTTTCTTATGTCCTCTT 840
Qy 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300
Db 841 ATGGCAGTCATTGGCAGCAGCTTCTTGTATTATTTCTCAAAGTAGCAGCATTTGTATATTT 900
Qy 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320
Db 901 CTGCTTTTCTTCTTATGGATTATCATCTGTATTATTTTGTCTTAAATGCTGACACCTCTT 960
Qy 321 PheLysLysSerIysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340
Db 961 TTTTAAAAATCAAAACATGCGGAATAGTTGAATTTTGTGCTGCTTGTGCTTGTGCTTGT 1020
Qy 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360
Db 1021 AITGGCTTATGATAATCTCATAGAAAGTTTTCCAAATCGTTAGTGTGGCTTTTCAGT 1080
Qy 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380
Db 1081 CCTTCTGTCAGCTACTTGTGTGATGTGTATGTCACAGGTGATGATTTAGAGATTTT 1140
Qy 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400
Db 1141 AATGAGGTGCTTCAATTTTCAAAATTTGACTGCAGGCCCATATCCCTAAATATTACAAT 1200
Qy 401 IleMetLeuThrIleAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420
Db 1201 ATCATGCTCACACTTAATAGTATATCTATGTCCTCTTGGCTGTCTATCTTCTGATCAAGTC 1260
Qy 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440
Db 1261 ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTATATTTTCTGAAGCTTTCATATGG 1320
Qy 441 SerLysSerIysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnLeuSer 460
Db 1321 TCAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGGCAATGTTAATGGAAATATTAGT 1380
Qy 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480
Db 1381 TTTAGTGAATATTATGAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCCATTAAGATT 1440
Qy 481 SerGlyIleGlnLysThrTyrArgLysLysGlyLysValGluAsnValGluAlaLeuArgAsnLeu 500
Db 1441 AGTGTATTTCAGAGACATACAGAAAGAGGTGAAATGTGAGGCTTTGAGAAATTTG 1500
Qy 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520
Db 1501 TCATTTGACATATATGAGGTCAGATTACTGCTTACTTTGGCCACAGTGGAAAGAGAAAG 1560
Qy 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540
Db 1561 AGTACATTGATGAATATCTTTGTGGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1620
Qy 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560
Db 1621 TATGGCAGACAGATCTCAGAAATAGATGAATGTTTGAAGCAAGAAATATGATGGCAAT 1680
Qy 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580
Db 1681 TGTCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAAGAAATTTATCAATTTTG 1740
Qy 581 AlaSerIleGlyGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600
Db 1741 GCTTCAATCAAAAGGGATACAGCCCAATATAATACAGAAAGTGCAGAAAGTTTACTTA 1800
Qy 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620
Db 1801 GATTTAGACATGCAGACTATCAAGATACCAAGCTTAAATAATTAAGTGGTGGTCAAAA 1860
Qy 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640
Db 1861 AGAAAGCTGTCATTAGGAATTTGCTGTTCTTGGAAACCCCAAGATATCTGCTGCTAGTAA 1920
```

995 LeuAenValThrGluThrIleGlnIleTrpSerThrProPheGlnGluIleThrAsp 1014  
Db TTAATGTGACTGAACCATCCAGATCTGGAGTACCCATCTCTTCAAGAAATTAAGTAT 3173  
1015 lleValPheIleGluLeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAl 1034  
Db ATAGTGTGTTTAAATTAAGTCTGATTTTCAAGAGCTTTGCTTGGAAATCATTTGACTGC 3233  
1034 aMetProProTyrPheAlaMetGluAenGluAenHisIleLysIleLysAlaTyrThrGl 1054  
Db AATGCCACCCTACTCTGCCATCGAAATGCGAGAAATCATAA----- 3275  
1054 nLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIl 1074  
Db -----GGTCTTTTCCCATCTGCATATTGGATTGGATTGGACAGCTGTGTGATAT 3322  
1074 eProLeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrGl 1094  
Db CCCCTTATTTTATCATCTTATTTTGTATGCTAGAGAGCTTATTTGGCATTTTCATTATGG 3382  
1094 yLeuTyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValPr 1114  
Db ATTATATTTTATCTGTAAGATCTCTGCTGGTGTGTGCTTTTCCCTATTGGTTATGTTCC 3442  
1114 oSerValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAenThrLy 1134  
Db ATCAGTTATCTGTTTACATATATGCTTCTTCCCTTTAAGAAAAATTTAAATATCCAA 3502  
1134 sGluPheTrpSerPheIleTyrSerValAlaLeuAlaCysIleAlaIleThrGluIl 1154  
Db AGAATTTTGTCTATTTATCTCTGTCGAGCTGTGCTTGTATTGGCAATCAGTGAAT 3562  
1154 eThrPhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIl 1174  
Db AACTTCTTATTCGGATACACAAATGCAATCTTCTTCAATATGCTTTTGTATCATCAT 3622  
1174 eProIleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAenVa 1194  
Db TCCAAATCTATCCACTCTAGTGTGCTGATTTCTTCAATAAGATTTCTTGAAGAATGT 3682  
1194 lArgLysAenValAspThrTyrAenProTrpAspArgLeuSerValAlaValIleSerPr 1214  
Db ACAGAAAAATGTGGACACCTATATTCATGGGATAGCTTTCAGTAGCTTTATATCGCC 3742  
1214 ofTyrLeuGlnCysValIleuTrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyGl 1234  
Db TTACCTGCACTGTGTACTGTGGATTTTCTCTTACAAATACTATGAGAAAAAATATGGAGG 3802  
1234 yArgSerIleArgLysAspProPhePheArgAenLeuSerThrLysSerLysAenArgLy 1254  
Db CAGATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTCAACGAGCTTAAAAATAGGA 3862  
1254 sLeuProGluProProAspAenGluAspGluAspGluValLysAlaGluArgLeuLy 1274  
Db GCTTCCAGAACCCACAGCAATAGGATGAAGATGAAGATGTCAAGAGCTGAAAGACTAAA 3922  
1274 sValLysGluLeuMetGlyCysGlnCysCysGluGluLysProSerIleMetValSerAs 1294  
Db GGTCAAGAGCTGATGTGGTGTGCGAGTGTGTGAGGAGAAACCATTCATATGCTGACGA 3982  
1294 nLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSerArgLysValLysLysVa 1314  
Db TTTGCATAAGAAATATCATGACAAGAAAGATTTCTTCTTCAAGAAAAAGTAAAGAAAGT 4042  
1314 lAlaThrLysTyrIleSerPheCysValLysLysGlyGluIleuGlyLeuLeuGlyPr 1334  
Db GGCAACTAAATACATCTCTTCTGTGTGAAAAAGGAGAGATCTTAGGACTATTGGGTCC 4102  
1334 oAsnGluValaGlyLysSerThrIleIleAenIleLeuValGlyAspIleGluProThrSe 1354  
Db AAAATGTTGCTGCGCAAAAGCACAAATTTATATTTCTGTTGGTGTATTTGAACCACTTC 4162

1354 rGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAspAspSerLeuLy 1374  
Db AGCCAGGATTTTATAGGAGATTTCTTCAGACAAGATGAAGATGATGATTCCTGAA 4222  
1374 sCysMetGlyTyrCysProGlnIleAenProLeuTrpProAspThrThrLeuGlnGluHi 1394  
Db GTGTATGGGTATCTGTCTCAGATAAACCTTTTGGCCGAGATCTACTACATTCGACGAACA 4282  
1394 sPheGluIleTyrGlyValaValLysGlyMetSerAlaSerAspMetLysGluValIleSe 1414  
Db TTTTGAATTTATGGAGCTGTCAAGGAATGATGCAAGTGCATGAAAGAAAGTCATTAAG 4342  
1414 rArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuPr 1434  
Db TCGAATAACACATGCACTTGTATTTAAAGAAACATCTTCAGAAGACTGTAAAGAACTACC 4402  
1434 oAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAenProGlnIleTh 1454  
Db TGCAGGAATCAACAGAAAGTGTGTGTTTCTCTAAGTATGCTAGGGAATCCTCAGATTAC 4462  
1454 rLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAl 1474  
Db TTTGCTAGATGAACCATCTACAGGTATGATCCCAAGGCCAAACACACATGTGGGAGC 4522  
1474 alIeArgThrAlaPheLysAenArgLysArgAlaAlaIleLeuThrThrHisTyrMetGl 1494  
Db AATTCGAACCTGCATTTTAAACACAGAAAGCGGCTGTATTCGACCACCTCATATATGA 4582  
1494 uGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIl 1514  
Db GGAGGCAGAGGCTGTCTGTATCGAGTAGCTATCATGCTGTCTGGGACAGTTAAGATGTAT 4642  
1514 eGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLe 1534  
Db CGGAACAGTAGTCAACATCTAAAGAGTAAATTTGGAAGAGCTACTTTTGGAAATTAAT 4702  
1534 uLysAspTrpIleGluAenLeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePh 1554  
Db GARGACTGATAGAAACCTTAGAAGTAGACCGCTTCAAGAGAAAAATTCAGTATATTT 4762  
1554 eProAsnAlaSerArgGlnGlnSerPheSerSerIleLeuAlaTyrLysIleProLysGl 1574  
Db CCCAAATGCAAGCCGTGAGAAAGTTTTCTTCTATTTTGGCTTATATAAATTCCTAAGGA 4822  
1574 uAspValGlnSerLeuSerGlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAl 1594  
Db AGATGTTCACTCCCTTTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4882  
1594 alIeGluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLy 1614  
Db CATTTGAAGATATATAGCTTTTCTCAAGCAACATTTGGAACAGCTTTTGTAGAATCTACTAA 4942  
1614 sGluGlnGluGluAspAenSerCysGlyThrLeuAenSerThrLeuTrpTrpGluAr 1634  
Db AGAACAGAGAGAGAGAGATATAGTGTGGNACTTTTAAACAGACACACTTTTGGTGGGAACG 5002  
1634 gThrGlnGluAspArgValValPhe 1642  
Db AACCAAGAGATAGATAGTATTT 5027

## RESULT 7

US-09-971-121-3  
Sequence 3, Application US/09971121  
Patent No. US20020111477A1  
GENERAL INFORMATION:  
APPLICANT: Hu, Yi  
APPLICANT: Nepomichy, Boris  
TITLE OF INVENTION: No. US20020111477A1 Human Transporter Proteins and Polynucleot  
TITLE OF INVENTION: Same  
FILE REFERENCE: LEX-0250-USA  
CURRENT APPLICATION NUMBER: US/09/971,121  
CURRENT FILING DATE: 2001-10-04  
PRIOR APPLICATION NUMBER: US 60/239,629



894 GCCTTTGGCTTCTGGGTTCTTATATACAGTTAAATTTCTTATGTCCCTTCTT 953  
 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300  
 954 ATGGCAGTCATTCGACAGCTCTCTTTGTTAATTCCTCAAGTAGCAGCAATTTGTATATTT 1013  
 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 1014 CTGCTTTTCTCTTATGGATATCACTGTATTTTGTCTTTAATGCTGACACCTCTT 1073  
 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 1074 TTTAAAAATCAAAACATGGGGAATAGTTGAATTTTGTGTACTGTGGCTTTTGGATTT 1133  
 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTyrLeuPheSer 360  
 1134 ATTGGCTTATGATTAATCTCATAGAAAGTTTCCAAATCTGTAGTGTGGCTTTTCAGT 1193  
 361 PropheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 1194 CCTTCTGTCACTGTACTTTTGTGATTTGTATTTGCACAGGTCATGCATTTAGAAGATTTT 1253  
 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 1254 AATGAAGGTGCTTCATTTCAAAATTTGACTCGAGGCCCATATCTCTCAATTAATTAACAAT 1313  
 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 1314 ATCATGCTCACACTTAATAGTATATTTCTATGTCTTGGCTGTCTATCTTGATCAAGTC 1373  
 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 1374 ATTCCAGGGGAATTTGGCTTAGGAGATCATCTTATATTTTCTGAAGCCTTCATATTGG 1433  
 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 1434 TCMAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGGCAATGTTAATGGAAATATTAGT 1493  
 461 PheSerGluIleIleGluProValSerSerGluPheValGlyIleGluAlaIleArgIle 480  
 1494 TTTAGTGAATATTATGAGCCAGTTTCTTCAGAAATTTGTAGGAAAAGAGCCATAAGAATT 1553  
 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
 1554 AGTGGTATTTCAGAGACATACAGAAAGAGGGTGAATGTGGAGCTTTGAGAAATTTG 1613  
 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 1614 TCATTTGACATATATGAGGTCAGATTACTTGCTTACTTTGGCCACAGTGGAAACAGGAAAG 1673  
 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540  
 1674 AGTACATTTGATGAATATCTTTGTGGACTTCGCCACCTTCATGTTGGTTTGCATCTATA 1733  
 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 1734 TATGGACACAGAGCTCTCAGAAATAGATGAAATGTTTGAAGCAAGAAAATGATTGGCATT 1793  
 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 1794 TGTCACAGTTAGATATACATTTGTATGTTTGTACAGTAGAAGAAAATTTATCAATTTTG 1853  
 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 1854 GCCTTCAATCAAGGGATACCAGCAACATATAATACAGAAGTGCAGAGGTTTTTACTA 1913  
 601 AspLeuAspMetClnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyClnLys 620  
 1914 GATTTAGACATGCAGACTATCAAGATACCAAGCTTAAATAATTAAGTGGTGGTCAAAA 1973  
 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640

1974 AGAAAGCTGTCAATTAGGAATTTGCTTCTTGGAAACCCAAAGATACCTGCTGTAGTAGAA 2033  
 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLeuLysTyrArg 660  
 2034 CCAACAGCTGGAATGAGACCCCTGTTCTCGACATATTTGATGGAATCTTTTAAATACAGA 2093  
 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 2094 AAAGCCAAATCGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATTTCTTGA 2153  
 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 2154 GATAGGAAGCTGTGATATCAAGAAATGCTGAAATGTGTGTTCTTCAATGTTCTCTC 2213  
 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 2214 AAAAGTAAATGGGGATCGGCTACCGCTGAGCATGTACATACACAAATATTGTGCACA 2273  
 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
 2274 GAATCTCTTTCTTCACTGGTTAAACACATATACCTGGAGCTACTTTATTACACAGAAAT 2333  
 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
 2334 GACCAACAATCTGTGTATAGCTTGCCTTCAAGACATGACACAAATTTTCAGGTTGTTT 2393  
 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
 2394 TCTGCCCTAGACAGTCAATCAAAATTTGGGTGTCATTTCTTATGTTGTTTCCATGAGCACT 2453  
 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
 2454 TTGGAGACGCTATTTTAAAGCTAGAGTTGAGCAGAAATTCACCAAGCAGATTATAGT 2513  
 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
 2514 GTATTTACTCAGCAGCCACTGGAGGAAGAAATGGAATCAAAATCTTTTGATGAATGGAA 2573  
 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
 2574 CAGAGCTTACTTATTTCTTCTGAACCCAGGCTGCTCTAGTAGCAGCATTGAGCCCTTGG 2633  
 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
 2634 AAACAACAGATGTATACATAGCAAGTTTCAATTTCTTTACCTTGAACCGTGAAGATAA 2693  
 861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
 2694 TCAGTGAAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTT 2753  
 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
 2754 TTGGTTTCATCACCTTTTAAATGCTGTGTTCCCATCAAACTTGTTCAGACTTATAT 2813  
 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
 2814 TTTCTAAACCTGGAGACAAACACATATAATACAAACAGTCTGCTCTTCTTCAAAATCT 2873  
 921 Ala-----AspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
 2874 GCTGGTGAGAGTCTNNGTGAGGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAGC 2933  
 935 GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
 2934 CAGAACTAATATGTTGAGCATTTAATGACAGTATGTATCCGTGGTCCCTCCATAGT 2993  
 955 AlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSer 974  
 2994 CGCGCTTTTAAATGTGATGCAATTCAGAAAGAGCATGTGTTTTGACAGCTGTGTTTCAACAGT 3053  
 975 ThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHis 994  
 3054 ACTATGGTTTATTTCTTTACCTATATTAGTGAATATCATTTAGTAGTAACACTACTTCTTATCAT 3113

```

Db 4050 AGCCAGGATATTTAGGAGATTTATCTTCAGAGCAAGTGAAGATGATTCACGTGAA 4109
Qy sCysMetGlyTyrCysProGlnIleAsnProLeuTyrProAspThrThrLeuGlnIle 1394
Db 4110 GTGTATGGTTACTGCTCCTCAGATAAACCCCTTTGTGGCCAGATACATTCAGGAACA 4169
Qy sPheGluLeuTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSe 1414
Db 4170 TTTTGAATTTATGGAGCTGTCAAGAGATGAGTGCAGTGCATGATGAAGAGTCTAAG 4229
Qy rArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnIleThrValLysLysLeuPr 1434
Db 4230 TCGAATACACATGCTGATTTAAAGAACATCTTCAGAGACTGTAAAGAACTACC 4289
Qy 1434 cAlaGlyIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleH 1454
Db 4290 TGCAGGAATCAACGAAAGTTGTGTTTGTCTTAAGTATGCTAGGGAATCCTCAGATTAC 4349
Qy 1454 rLeuLeuAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAl 1474
Db 4350 TTTTGTAGATGAACCATCTCAGGTATGGATCCCAAGCCCAACAGCACATGTGGCGAGC 4409
Qy 1474 aileArgThrAlaPheLysAsnArgLysAlaAlaIleLeuThrThrHisTyrMetGl 1494
Db 4410 AATTCGAACTGCATTTAAACACAGAAAGCGGCTGCTATTCTGACCACTCATATATGA 4469
Qy 1494 uGluAlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIl 1514
Db 4470 GGAGCGAGAGCTGCTGTGATCGATGAGTATCATGCTGTGCTGGCGAGTTAAGATGTAT 4529
Qy 1514 eGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLe 1534
Db 4530 CGGAACAGTACACATCTAAAGAGTAAATTTGGAAAGGCTACTTTTGGAAATTAAT 4589
Qy 1534 uIleAspTyrIleGlnAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePh 1554
Db 4590 GAAGACTGGATAGAAACCTTAGAAGTAGACCGCTTCAAGAGAAATTCAGTATATTT 4649
Qy 1554 eProAsnAlaSerArgGlnGlnSerPheSerSerIleLeuAlaTyrLysIleProLysGl 1574
Db 4650 CCCAAATGCAAGCGCTCAGGAAGTATTTCTCTATTTGGCTTTATAAATTCCTAAGGA 4709
Qy 1574 uAspValGlnSerLeuSerGlnSerPheLysLeuGluAlaLysHisAlaPheAl 1594
Db 4710 AGATGTTCAAGCTCTTCAACATCTTTTAAAGCTGGAGAGCTAAACATGCTTTTGC 4769
Qy 1594 aileGluGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrTy 1614
Db 4770 CATTGAAGAAATATAGCTTTTCTCAAGCAACATTTGGAAACAGTTTTGTAGAACTCACTAA 4829
Qy 1614 eGluGlnGluGluAspAsnSerCysGlyThrLeuAsnSerThrLeuTyrTyrGluAr 1634
Db 4830 AGAACAGAGGAGGAAGATATAGTTGTGGAACCTTTAAACAGCACACTTTGTGGGAACG 4889
Qy 1634 gThrGlnGluAspArgValPhe 1642
Db 4890 AACACAGAGATAGATGATTT 4914

```

RESULT 6

```

US-10-090-458-1
; Sequence 1, Application US/10090458
; Publication No. US20020123107A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hongyun
; APPLICANT: Killinski, Ligia
; APPLICANT: Le Bihan, Stephane
; TITLE OF INVENTION: NOVEL ASCAS TRANSPORTER AND USES THEREOF
; FILE REFERENCE: 100103.403
; CURRENT APPLICATION NUMBER: US/10/090.458
; CURRENT FILING DATE: 2002-03-01
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 1
; LENGTH: 5463
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 25, 2888, 2889
; OTHER INFORMATION: n = A, T, C or G
US-10-090-458-1

Alignment Scores:
Pred. No.: 0 Length: 5463
Score: 8309.00 Matches: 1629
Percent Similarity: 98.85% Conservative: 1
Best Local Similarity: 98.79% Mismatches: 2
Query Match: 98.61% Indels: 18
DB: 13 Gaps: 2

US-10-090-458-5 (1-1642) x US-10-090-458-1 (1-5463)

Qy 1 MetSerThrAlaIleArgGluValGlyValTyrArgGlnThrArgThrLeuLeuLys 20
Db 114 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGGAGACAGACCAGAACACTTCTACTGAAG 173
Qy 21 AsnTyrLeuLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu 40
Db 174 AATTCTTATTAATGCGAACCACCAAAAGAGTAGTGTTCAGAAATTCCTTTTCCACTA 233
Qy 41 PhePheLeuPheTyrLeuLeuLeuSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 234 TTTTCTTATTTTGGTTAATTAATTAGCATGATGATGCCATCCAAATAAGAAATATGAAGA 293
Qy 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80
Db 294 GTCCCTAATATAGAACTCAATCCTATGGACAGTTTACTTCTTCTTATCTTAACTTCTTGA 353
Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
Db 354 TATATCCAGTGACTAATATTACAAGCAGCATCATGCAAGAGTGTCTACTGATCATCTA 413
Qy 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
Db 414 CTTGATGTCATAATTTACTGAGAAATATACAAATGAAAGAAATGTTTAACTCCAGTCTC 473
Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
Db 474 TCTAAGCCGAGCAACTTTGTAGGTGTGGTTTCAAAGACTCCATGCTCTATGAATTCGT 533
Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
Db 534 TTTTCTCTGATATGATTCAGTATCTTCTATTATATGGATTCAAGAGCTGGCTGTTC 593
Qy 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180
Db 594 AAATCATGTGAGGCTGCTCAGTACTGGTCTCAGGTTTTCACAGTTTTCACAGCATCCATA 653
Qy 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGluSerThr 200
Db 654 GATGCTGCCATATATACATTTGAGACCAATGTTTCTCTTGGAGAGGAGCTGGAGTCACT 713
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
Db 714 AAAGCTGTTTATTATGGAGAAACCTGCTGTGTGTAGAAATAGATACCTTTTCCCGAGGAGTA 773
Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
Db 774 ATTTTATATACCTAGTTATAGCAATTTTCCCTTTTGGATCTTTTGGCAATTCATATC 833
Qy 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
Db 834 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGGACTTCATGATACT 893
Qy 261 AlaPheTyrLeuSerTyrValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280

```

QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuIleValArg 560  
Db 1921 CCAACAGCTGGAATGGACCCCTGCTCTCGACATATTGTATGGAATCTTTTAAATACAGA 1980  
QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 580  
Db 1981 AAGCCATCGGGTGACAGTGTTCAGTACTCATTTTCATGGATGAAGCTGACATCTTGGCA 2040  
QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
Db 2041 GATAGGAAGCTGTGATATCACAGGAATGCTGAAATGTGTGTTCTTCAATGTTCCCTC 2100  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
Db 2101 AAAAGTAATGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACA 2160  
QY 721 GluSerIleSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
Db 2161 GAATCTCTTCTCACTGGTTAAACAACATATACCTGGAGCTACTTTATTACACAGAA 2220  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
Db 2221 GACCAACAACCTGTGTATAGCTTGCCTTTCAAGGACATGGACAAATTTTCAGGTTTGT 2280  
QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
Db 2281 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCACTTTCTATGTGTGTTCCATGACGACT 2340  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
Db 2341 TTGGAAGACGCTATTTTAAAGCTAGAAAGTTGAAGCAGAAATTTGACCAAGCAGATTATAGT 2400  
QY 801 ValPheThrGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
Db 2401 GTATTACTCAGCAGCCATCGAGGAGAAATGGATTCMAAATCTTTTGATGAAATGGAA 2460  
QY 821 GlnSerLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
Db 2461 CAGAGCTTACTTATCTTCTGAAACCAAGGCTGCTAGTGAGCACCACCTGAGCCTTTGG 2520  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
Db 2521 AAACACAGATGTATACAAATAGCAAGTTTCACTTTCTTACCTTGAACGGTGAAGTAAA 2580  
QY 861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
Db 2581 TCAGTCAGATCAGTGTGCTTCTGCTTTTAAATTTTTCACAGTTTCAGATTTTATGTTT 2640  
QY 881 LeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
Db 2641 TTGGTTTCATCTCTTTTAAATGCTGTGGTTCCTCATCAAACTTGTTCAGACTTATAT 2700  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
Db 2701 TTTCTAAACCTGGAGACAAACCATAAATACAAAACAGTCTGCTCTTCTCAAAATCT 2760  
QY 921 Ala-----AspSerAspIleSerAspLeuIleSerPhePheThrSer 934  
Db 2761 GCTGGTGAGAGGTGNGGTGAAGACTCAGATATCAGTGATCTTATTAGCTTTTTCACAAGC 2820  
QY 935 GlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSer 954  
Db 2821 CAGACATAATGGTGACCATGATTAATACAGTGTACTATGATCCGTGGCTCCCATAGT 2880  
QY 955 AlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSer 974  
Db 2881 CGCGCTTTAAATGTGATCAITTCAGAAAGGACTATGTTTTTGCAGCTGTTTTCAACAGT 2940  
QY 975 ThrMetValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHis 994  
Db 2941 ACTATGGTTTATCTTTACCTATATATAGTAATATCATTAGTAACACTACTATCTTTATCAT 3000

QY 995 LeuAsnValThrGluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAsp 1014  
Db 3001 TTAATGTGACTGAAACCCATCCAGATCTGGAGTACCCCATCTTCTTCAAGAAATTTACTGAT 3060  
QY 1015 IleValPheTyrIleGluLeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAl 1034  
Db 3061 ATAGTTTTTAAATTTGAGCTGTATTTTCAAGCAGCTTTGCTTGGAAATCATTTGAACTGC 3120  
QY 1034 aMetProProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrG1 1054  
Db 3121 AATGCCACCTTACTCTGCCATGGAAATGCAGAAATCAATA----- 3162  
QY 1054 nLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspI1 1074  
Db 3163 -----CGTCTTTTGCATCTGCATATTTGATTTGGACAGCTGTTGTTGATAT 3209  
QY 1074 eProLeuPhePheIleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrG1 1094  
Db 3210 CCCCTTATTTTATCATCTTATTTTGTGCTAGGAGCTTATTTGGCATTTCTATTATGG 3269  
QY 1094 YLeuTyrPheTyrThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValPr 1114  
Db 3270 ATTATATTTTATCTGTAAGTTTCTTGTGTGGTTTTTTTGGCTTATTGCTTATGTTCC 3329  
QY 1114 oSerValIleLeuPheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLys 1134  
Db 3330 ATCAGTTATTTCTGTTCACTTATTTGCTTCTTTCACCTTTTAAAGAAATTTTAAATACCAA 3389  
QY 1134 sGluPheTrpSerPheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluI1 1154  
Db 3390 AGAATTTTGTGTCATTTATCTTATCTGTGGCAGCGTTGGCTTGTATTGCAATCACTGAAAT 3449  
QY 1154 eThrPhePheMetGlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleI1 1174  
Db 3450 AACTTTCTTATGGGATACACAATTCGACTTATCTTCTTATTATGCTTTTGTATCATCAT 3509  
QY 1174 eProIleTyrProLeuLeuGlyCysLeuIleSerPheIleLysIleSerTyrLysAsnVa 1194  
Db 3510 TCCAAATCAATCCACTTCTAGTTTGCCTGATTTCTTTCATAAAGATTTCTTGAAGAATGT 3569  
QY 1194 LArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSerPr 1214  
Db 3570 ACGAAAAATGTGGACCCATTAATCCATGGGATAGGCTTTCAGTAGCTGTTTATATCGCC 3629  
QY 1214 oTyrLeuGlnCysValIleTrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyG1 1234  
Db 3630 TTACCTGCAGTGTGACTGTGGATTTCTCTTACATACTATGAGAAAAATATGGAGG 3689  
QY 1234 YArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLys 1254  
Db 3690 CAGATCAATAAGAAAAAGATCCCTTTTTCAGAAACCTTTTCAACGAGCTCTAAAAATAGGAA 3749  
QY 1254 sLeuProGluProProAspAsnGluAspGluAspValLysAlaGluArgLeuLys 1274  
Db 3750 GCTTCCAGAACCCACACAGCAATGAGGATGAAGATGAAGATGTCAGAAAGTGAAGACTAAA 3809  
QY 1274 sValLysGluLeuMetGlyCysGlnCysCysGluGluLysProSerIleMetValSerAs 1294  
Db 3810 GGTCAAGAGCTGATGGTTTGCAGTGTGTTGTGAGGAGAAACCATCCATTTATGGTCAGCAA 3869  
QY 1294 nLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSerArgLysValLysLysVa 1314  
Db 3870 TTTGCAATAAGAAATATGATGACAAAGAAATTTTCTTCTTTTCAAGAAAAATGAAGAAAGT 3929  
QY 1314 lAlaThrLysTyrIleSerPheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyPr 1334  
Db 3930 GGCACATAATACATCTCTTTCTGTGTGAAAAAGAGAGAGATCTTAGGACTTATTTGGTCC 3989  
QY 1334 oAsnGlyAlaGlyLysSerThrIleIleAsnIleLeuValGlyAspIleGluProThrSe 1354  
Db 3990 AAATGGTGTGCAAAAGACAAATTTATTAATTTCTGTTGTTGTTGATATTATGAACCACTTC 4049  
QY 1354 rGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGluAspAspSerLeuLys 1374

```

; LENGTH: 4917
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2775..2776
; OTHER INFORMATION: n = A,T,C or G
US-10-090-458-3

Alignment Scores:
Pred. No.: 0 Length: 4917
Score: 8309.00 Matches: 1629
Percent Similarity: 98.85% Conservatives: 1
Best Local Similarity: 98.79% Mismatches: 2
Query Match: 98.61% Indels: 18
DB: 13 Gaps: 2

US-10-090-458-5 (1-1642) x US-10-090-458-3 (1-4917)

Qy 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20
Db 1 ATGTCCACTGCAATTAGGAGGTAGAGTTTGAGACAGACAGACAGCAACTTCTACTGAAG 60
Qy 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGluIleLeuPheProLeu 40
Db 61 AATTACTTAATAATGCGAAGCAACCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 120
Qy 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 121 TTTTITTTTATTTGGTTAATAATAATAGCATGATGCATCCAAATAAGAAATATAGAGAA 180
Qy 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80
Db 181 GTGCCTAATATAGAACTCAATCTATGGACAAGTTTACTTCTTCTTAATCTAATCTTGA 240
Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
Db 241 TATACTCCAGTCACTAATATTAACAGCAGCATCATGCAAGAAAGTCTCTACTGATCATCTA 300
Qy 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
Db 301 CCTGATGCTAATAATTAAGAGATATACAAATGAAAGAAATGTTAAACATCCAGTCTC 360
Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
Db 361 TCTAAGCCGAGCAACTTTCTAGTGTGGTTTCAAAGACTCCATGCTCCTATGAACTTCGT 420
Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
Db 421 TTTTTCCTGATATGATTCAGTAATCTTCTATTTATATGATTCAGAGCTGGCTGTCA 480
Qy 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180
Db 481 AAATCATGTGAGGCTGCTCAGTACTGCTCTCAGGTTTCAACAGTTTTCACAGCATCCATA 540
Qy 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200
Db 541 GATGCTGCCATATATACAGTTGAGACCAATGTTTCTTTTGAAGAGAGCTGGAGTCAACT 600
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
Db 601 AAAGCTGTTATTATGGAGAAACTGCTGTGTAGAAATAGATACCTTTCCCGAGAGAGTA 660
Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
Db 661 ATTTTAAATATACCTAGTTATAGCATTTTCACCTTTTGGATATCTTTTGGCAATTCATATC 720
Qy 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
Db 721 GTAGCAGAAAAAGAAAAAATAAAGAAATTTTAAAGATAATGGACTTCATGATACT 780
Qy 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280

```

```

Db 781 GCCTTTTGGCTTTCCTGGGTTCTTCTATATACAAAGTTTAAATTTTCTTATGTCCTCTCT 840
Qy 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300
Db 841 ATGGCAGTCATTGGCGACAGCTTCTTTGTTATTCTCTCAAAGTAGCAGCATTTGTGATAATT 900
Qy 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320
Db 901 CTGCTTTTTCCTTTATGGATTATCATCTGATTATTTTTCCTTTTATGCTGACACTCTTT 960
Qy 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340
Db 961 TTTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTGTTTACTGTGGCTTTTGGATT 1020
Qy 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360
Db 1021 ATTGGCCTTATGATAATCTCTCATAGAAAGTTTCCCAAATCGTTAGTGGGCTTTTCAGT 1080
Qy 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380
Db 1081 CCTTTCTGTACCTGTAATTTTGTGATTGTTGACAGTCATGTCATTTAGAGATTTT 1140
Qy 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400
Db 1141 AATGAAGGTGCTTCATTTTCAAATTTGACTGAGGCCCATATCCCTCTAATTTATCAATT 1200
Qy 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420
Db 1201 ATCATGCTCACCTTAATAGTATATTTCTATGCTTCTGCTGCTATCTTGTATCAAGTC 1260
Qy 421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440
Db 1261 ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTTCATATTG 1320
Qy 441 SerLysSerIleArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460
Db 1321 TCAAAGACAAAGAAATTTATGAGAGTTATCGAGGCAATGTTAATGGAATATTAGT 1380
Qy 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480
Db 1381 TTTAGTGAAATTTAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCCATAGAAT 1440
Qy 481 SerGlyIleGlnLysThrTyrArgLysGlyGluAsnValGluAlaLeuArgAsnLeu 500
Db 1441 AGTGGTATTGAGAGACATACAGAAAGAGGGTGAATTTGGAGGCTTTGAGAAATTTG 1500
Qy 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520
Db 1501 TCATTTGACATATATGAGGTCAGATTCTGCCTTACTTTGGCCACAGTGGAAACAGGAAAG 1560
Qy 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540
Db 1561 AGTACATTGATGAATATTTCTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1620
Qy 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560
Db 1621 TATGGACACAGCTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATTGGCAATT 1680
Qy 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580
Db 1681 TGTCACAGTAGATATACACTTTGATGTTTGGACAGTAGAAGAAATTTATCAATTTTG 1740
Qy 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600
Db 1741 GCTTCAATCAAGGGATACCAGCCCAACATATAATACAGAGAGTGAGAGGTTTACTA 1800
Qy 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620
Db 1801 GATTAGACATGACAGACTATCAAGATAACCAAGCTAAAAAATAAGTGGTGGTCAAAAA 1860
Qy 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640
Db 1861 AGAAAGCTGTCATTAGGAATTTGCTTTTGGGAACCCAAAGATATCTGCTGATGATA 1920

```

Db 3255 ATCCAGATCTGGAGTACCCCTTTCTTCAAGAAATTAAGTATAGTTTTTAAATTTGAG 3314  
Qy LeuTyrPheGlnAlaAlaLeuLeuGlyTllelleValThrAlaMetProProTyrPheAla 1040  
Db 3315 CTGTAATTTCAAGCAGCTTTCCTTGGAAATCAATGTTACTGCAATGCCACCTTACTTTGCC 3374  
Qy MetGluAenAlaGluAenHisLysLysLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
Db 3375 ATGGAAATGCAGAGAAATCAATAGATCAAAAGCTTAYACTCAACTTAACTTTTCAGGTCCT 3434  
Qy LeuProSerAlaTyrTrrPileGlnAlaValValAspPileProLeuPhePheLlelle 1080  
Db 3435 TTGCCATCTGCATATATGGAATGGCAAGCTTGTGTGATATCCCTTATTTTATCATTT 3494  
Qy LeuLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
Db 3495 CTATTTTGTAGCTAGGAGCTTATGGCAATTCATATGGAATATATTTTATACGTA 3554  
Qy LysPheLeuAlaValValPheCysLeuLleGlyTyrValProSerValIleLeuPheThr 1120  
Db 3555 AAGTTCCCTTGTGTGGTATTTTGTGCTTATTTGGTATGTTCCATCAGTTATTTCTGTTCACT 3614  
Qy TyrIleAlaSerPheThrPheLysLysLysLeuAenThrLysGluPheTrrPhePheLle 1140  
Db 3615 TATATGCTCTTTCACCTTTAAGAAATTTTAAATACCAAGAAATTTTGTCTATTTATC 3674  
Qy TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
Db 3675 TATTCGTGGCAGCGTGGCTGTATTGCAATCACTCAATAAATCTTCTTTATGGGATAC 3734  
Qy ThrIleAlaThrIleLeuHisTyrAlaPheCysIlellelleProIleTyrProLeuLeu 1180  
Db 3735 ACAATGCAACTATTTCTTCAATATGCTTTTGTATCATCATTCATTCATTCACCTTCTA 3794  
Qy GlyCysLeuIleSerPheIleLysLysLysSerTrrPlysAenValArgLysAenValAspThr 1200  
Db 3795 GGTTCCTGATTTCTTTCATAAAGATTTCTTGGAGAAATGTACGAAATAATGTGCACACC 3854  
Qy TyrAenProTrrPheArgLeuSerValAlaValIleSerProTrrPheGlnCysValLeu 1220  
Db 3855 TATAATCATCGGATAGCTTTCAGTACCTGTTATATCGCCTTACCTCGCATGTGTACTG 3914  
Qy TrrPheLeuLeuGlnTyrTrrGluLysLysTrrGlyArgSerIleArgLysAsp 1240  
Db 3915 TGGATTTTCTTCTTACATATCTAGAAATAATATGGAGCAGATCATTAAGAAAGAT 3974  
Qy ProPhePheArgAenLeuSerThrLysSerLysAenArgLysLeuProGluProProAsp 1260  
Db 3975 CCTTTTTCAGAAACCTTTCAACGAAGTCAAAAATAGGAAGCTTCCAGAACCAACCCAGAC 4034  
Qy AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db 4035 AATGAGGATGAAGATGAAGATGTCAAAGCTGAAGACTTAAGGTCAAGAGCTGATGGGT 4094  
Qy CysGlnCysCysGluGluLysProSerIleMetValSerAenLeuHisLysGluTyrAsp 1300  
Db 4095 TGCCAGTGTCTGAGGAGAACCATTCATTTATGCTCAGCAATTTGCAATAAGAAATATGAT 4154  
Qy AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTrrLysSer 1320  
Db 4155 GACAAGAAAGATTTCTTCTTCAAGAAAGATGAAGAAATGGCACTAAATAATCATCTCT 4214  
Qy PheCysValLysLysGlyGluLeuLeuGlyLeuLeuGlyProAenGlyAlaGlyLysSer 1340  
Db 4215 TTCTGTGAAAAAAGAGAGATCTTAGGACTATTGGTCCAAATGGTGTGCGCAAGAGC 4274  
Qy ThrIlelleAenIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db 4275 ACAATTATTAATATCTGGTGGTGCATATTAACCAACTTCAGCCAGGATTTTATGGA 4334  
Qy AspTrrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTrrCysPro 1380

Db 4335 GATTATTTCTCAGACAGAGTGAAGATGATGATTTCACTGAAGTGTATGGTTACTGTCT 4394  
Qy GlnIleAenProLeuTrrProAspThrThrLeuGlnGluHisPheGluIleTrrGlyAla 1400  
Db 4395 CAGATAAACCCCTTTGTGGCCAGATACATTTGCGAGAACATTTTGAATTTTATGGAGCT 4454  
Qy ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db 4455 GTCAGAGGAATGAGTGAAGTGCATGAAGAAGTCATAAGTCGAATAACACATGCACCTT 4514  
Qy AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyLleLysArgLys 1440  
Db 4515 GATTTAAAGAACATCTTTCAGAGAGCTGTAAAGAACTACCTGCGAGGAATCAACAGCAAG 4574  
Qy LeuCysPheAlaLeuSerMetLeuGlyAenProGlnIleThrLeuLeuAspGluProSer 1460  
Db 4575 TTGTGTTTGTCTTAAGTATGCTAGGAATCTCTCAGATTAATTTTGTAGTAGAACCACT 4634  
Qy ThrGlyMetAspProLysAlaLysGlnHisMetTrrPheAlaIleArgThrAlaPheLys 1480  
Db 4635 ACAGGTATGGATCCCAAGCCAAACAGCACATGTGGCGAGCAATTCGAACATGCATTTAA 4694  
Qy AsnArgLysArgAlaAlaIleLeuThrThrHisTrrMetGluGluAlaGluAlaValCys 1500  
Db 4695 AACAGAAAGCGGCTGCTATTTCTGACCTACTATATGGAGGAGCGAGGCTGTCTGT 4754  
Qy AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
Db 4755 GATCGAGTAGCTATCATGTGTCTGGCAGTTAAGATGTATCGGAACAGTACCAATCTA 4814  
Qy LysSerLysPheGlyGlyGlyTrrPheLeuGluIleLysLeuLysAspTrrIleGluAen 1540  
Db 4815 AAGAGTAAATTTGGAAAGGCTACTTTTTTGGAAATTAATTTGAAGGACTGGATAGAAAC 4874  
Qy LeuGluValAspArgLeuGlnArgGluIleGlnTrrIlePheProAenAlaSerArgGln 1560  
Db 4875 CTAGAGTAGACCGCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCCGTGAG 4934  
Qy GluSerPheSerSerIleLeuAlaTrrLysLysLysProLysGluAspValGlnSerLeuSer 1580  
Db 4935 GAAAGTTTTTCTTATTTTGGCTTATAAAATTCCTAAGGAAGATGTTTCAGTCCCTTTCA 4994  
Qy GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTrrSerPhe 1600  
Db 4995 CAATCTTTTTTAAAGCTGGAGAGACTAAACATGCTTTTGCNTTGAAGATATAGCTTT 5054  
Qy SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
Db 5055 TCTCAAGCAACATTCGAACAGGTTTTTGTAGAACTCACTAAAGAACCAAGAGAGAGAT 5114  
Qy AsnSerCysGlyThrLeuAenSerThrLeuTrrTrrGluArgThrGlnGluAspArgVal 1640  
Db 5115 AATAGTTGGGAACCTTTAAACAGCACACTTTTGGTGGGAACGAACACAAGAGATAGATGA 5174  
Qy ValPhe 1642  
Db 5175 GTATTT 5180

## RESULT 5

US-10-090-458-3  
; Sequence 3, Application US/10090458  
; Publication NO. US20020123107A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Hongyun  
; APPLICANT: Killinski, Ligia  
; APPLICANT: Le Bihan, Stephane  
; TITLE OF INVENTION: NOVEL ABCA5 TRANSPORTER AND USES THEREOF  
; FILE REFERENCE: 100103.403  
; CURRENT APPLICATION NUMBER: US/10/090.458  
; CURRENT FILING DATE: 2002-03-01  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3

QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerIleValIlePhe 300  
 Db 1095 ATGGCAGTCATTGCCAGACGCTCTCTTTGATATTTCTCTCAAGTAGCAGCATTTGTATTT 1154  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 Db 1155 CTGCTTTTTCTCTTATGGATTATCATCTGTATTTTGTGTTTAAATGCTGACACCTCTT 1214  
 QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 Db 1215 TTTAAAAAATCAAAACATGCGGAATAGTTGAAATTTTTTTGTTACTGTGGCTTTTGGATTT 1274  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValThrLeuPheSer 360  
 Db 1275 ATTGCCATTATGATAATCTCATAGAAAGTTTTCCTCAATCGTTAGTGTGGCTTTTCAGT 1334  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
 Db 1335 CCTTTCTGTCACGTACTTTTGTGATTGTGATTGTGATTTGCACAGGTCATGCATTTAGAGATTTT 1394  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 Db 1395 AATGAGGTGCTTCATTTTCAAAATTTGACTGCGAGGCCATATCTCTAAATTTACAAAT 1454  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 Db 1455 ATCATGCTCACACTTAATAGTATATTTCTATGTCCTCTGGCTGTCTATCTTTGATCAGATC 1514  
 QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 Db 1515 ATTCAGGGGAAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTGG 1574  
 QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnLysIleSer 460  
 Db 1575 TCAAGAGCAAAAGAAATATGAGGAGTTATCAGAGGGCAATGTTAAATGGAAATATTAGT 1634  
 QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLeuAlaIleArgIle 480  
 Db 1635 TTTAGTGAATATTTCAGCCAGTTCTTCAGAAATTTGTAGGAAAGAACCCATAAGAATT 1694  
 QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
 Db 1695 AGTGGTATTTCAGAAACATACAGAAAGAGGGTGAAATGTGGAGCTTTTGAAATTTG 1754  
 QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 Db 1755 TCATTTGACATATATGAGGGTCAGATTACTGCTTACTTTGCCACAGTGCACAGGAAAG 1814  
 QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
 Db 1815 AGTACATTGATGAATATTTCTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1874  
 QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 Db 1875 ATGGACACAGATCTCAGAAATAGATGAATGTTTGAAGCAGAGAAATGATTTGGCATT 1934  
 QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 Db 1935 TGTCCACAGTTAGATATACATTTGATGTTTGTGACAGTAGAAGAAATTTATCAATTTTG 1994  
 QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 Db 1995 GCTTCAATCAAGGGATACCAGCCCAATATATATCAAGAGTGCAGAGGTTTACTATA 2054  
 QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 Db 2055 GATTTAGACATGCAGACTATCAAGATAACCAAGCTAAAAATTTAAGTGGTGTCAAAA 2114  
 QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuAspGlu 640  
 Db 2115 AGAAGCTGTCTTAGGAATTTGTTCTTTGGGAACCCCAAGATACTGCTGTAGATGAA 2174

QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTyrAsnLeuLeuLysTyrArg 660  
 Db 2175 CCAACAGCTGGAAATGGACCCCTGTTCTGACATATTTGATGAAATCTTTTAAATACAGA 2234  
 QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 Db 2235 AAAGCCATNTGGGTGACAGTGTTCAGTACTCATTTTCATGATGAAGCTGACATTTCTGCA 2294  
 QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 Db 2295 GATAGAAAAGCTGTGATATCACAAAGAAATGCTGAAATGTGTGGTCTCTCAATGTCCTC 2354  
 QY 701 LysSerLysIleGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 Db 2355 AAAGTAATNTGGGGATCGCTACCGCCCTGACGATGATACATAGACAAATATTGTCCACA 2414  
 QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
 Db 2415 GAATCTCTTTCTTCACTGGTTAAACAACATATACCTGGAGCTACTTTATTAACACAGAA 2474  
 QY 741 AspGlnGluLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
 Db 2475 GACCAACAACCTTGTATAGCTTGCCTTTCAGGACATGACACAAATTTTCAGGTTGTTT 2534  
 QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
 Db 2535 TCTGCCCTAGACAGTCATTCAAAATTTGGGTGGCAATTTCTTATGGGTTTCCATGACGACT 2594  
 QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
 Db 2595 TTGGAGACGTATTTTAAAGCTTAGAAGTTGAAGCAAAATTTGACCAACAGCATATAGT 2654  
 QY 801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
 Db 2655 GTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAATCTTTTATGAAATGGAA 2714  
 QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
 Db 2715 CAGAGCTTACTTATTTCTGAAACCAAGGCTCTCTAGTGAGCACCATGAGCCCTTGG 2774  
 QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
 Db 2775 AAACAACAGATGATACATAGCAAGTTTCATTTCTTACCTTGAACCGTGAAGATAA 2834  
 QY 861 SerValArgSerValLeuLeuLeuLeuLeuLeuPhePheThrValGlnIlePheMetPhe 880  
 Db 2835 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTCACAGTTTTCAGATTTTATGTTT 2894  
 QY 881 LeuValHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
 Db 2895 TTGGTTTCATCTCTTTTAAATATGCTGTGTTCCCATCAAACTTGTTCAGACTTATAT 2954  
 QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGlnAsnSer 920  
 Db 2955 TTTCTAAACCTGGAGACAAACACATAAATACAAACAAAGCTGCTTCTTCAAAATTTCT 3014  
 QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
 Db 3015 GCTGACTCAGATATCAGTGATCTTATGATCTTTTTCACAGCCAGACATATATGGTGACG 3074  
 QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
 Db 3075 ATGATTAATGACAGTACTATGATCGTGGCTCCCATAGTGGCGCTTTAAATGTGRTG 3134  
 QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrThrMetValTyrSerLeu 980  
 Db 3135 CATTCAAGAAAGGACTATGTTTTCAGAGCTGTTTTCACACAGTACTATGGTTTATCTTTA 3194  
 QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
 Db 3195 CCTATATAGTAGAATACATAGTACTACTATCTTTATCATTTAAATGTGACTGAAACC 3254  
 QY 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020

```
Db 4081 GATTATCTTCAGACCAAGTGAAGATGATGATCTCACTGAAGTGTATGGTTACTGTCT 4140
Qy 1381 GlnileAsnProLeuThrProLeuThrLeuGlnGluHisPheGluLeuIleValGlyAla 1400
Db 4141 CAGATAAACCCTTTGGCCAGACTACTACATTCAGAGAACATTTGCAATTTATGAGCT 4200
Qy 1401 ValLysGlyMetSerAlaSerMetLysGluValLysLeuValLysArgIleThrHisAlaLeu 1420
Db 4201 GTCAAGGAATGAGTGCAGTGCATGAAGAAGTCATAAGTCGAATCAACATCACTT 4260
Qy 1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGluLysArgLys 1440
Db 4261 GATTTAAAGACATCTTCAGAGACTGTAAAGAACTTACCTGCAGGAATCAAGCAAG 4320
Qy 1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460
Db 4321 TTGTGTTTTGCTCTAAGTATGCTAGGGAATCCTCAGATTACTTTGCTAGATGAACATCT 4380
Qy 1461 ThrGlyMetAspProLysAlaLysGlnHisMetTyrArgAlaIleArgThrAlaPheLys 1480
Db 4381 ACAGGTATGATCCCAAGCCAAACAGACATGTGGCGAGCAATTCGAATGCATTTAA 4440
Qy 1481 AsnArgLysArgAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCys 1500
Db 4441 AACAGAAAGCGGCTCTATCTGACCACTCATATATGAGAGGCGAGGCTGTCTGT 4500
Qy 1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520
Db 4501 GATCGAGTAGCTATCATGTGTCTGGCGAGTTAAGATGTATCGGAACAGTACACATCTA 4560
Qy 1521 LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTyrIleGluAsn 1540
Db 4561 AAGAGTAATTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGACTGGATAGAAAAC 4620
Qy 1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560
Db 4621 CTAGAAGTAGCGCCCTCAAAGAGAAATTCAGTATATTTTCCCAATGCAAGCGCTCAG 4680
Qy 1561 GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580
Db 4681 GAAAGTTTTCTTCTATTGCTTATAAATTCCTAAGGAAGATGTTCAAGTCCCTTTCA 4740
Qy 1581 GlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600
Db 4741 CAATCTTTTTTAAGCTGAAGAAGCTAAACATGCTTTTGGCAATGCAAGAATATAGCTTT 4800
Qy 1601 SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620
Db 4801 TCTCAAGCAACATTGGAACAGGTTTTTGTAGAACTCACTAAAGAACAGAGAGGAAGAT 4860
Qy 1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyrTyrGluArgThrGlnGluAspArgVal 1640
Db 4861 AATAGTTGTGGAACCTTAAACAGCACACTTTGGTGGGAACGAACAACAAGATAGAGTA 4920
Qy 1641 ValPhe 1642
Db 4921 GATATT 4926
```

## RESULT 4

```
US-09-971-121-5
; Sequence 5, Application US/09971121
; Patent No. US20020111477A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Repomnichy, Boris
; TITLE OF INVENTION: No. US20020111477A1 Human Transporter Proteins and Polynucleotides
; TITLE OF INVENTION: Same
; FILE REFERENCE: LEX-0250-USA
; CURRENT APPLICATION NUMBER: US/09/971,121
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: US 60/239,629
; PRIOR FILING DATE: 2000-10-10
```

```
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 5262
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-971-121-5

Alignment Scores:
Pred. No.: 0
Score: 8395.00
Percent Similarity: 99.70%
Best Local Similarity: 99.70%
Mismatch: 5
Indels: 0
Query Match: 99.63%
Gaps: 9

US-10-090-458-5 (1-1642) x US-09-971-121-5 (1-5262)

Qy 1 MetSerThrAlaIleAlaArgGluValGlyValTyrArgGlnThrArgThrLeuLeuLys 20
Db 255 ATGTCCACTGCAATTAGGGAGGTAGGAGTTTGGAGACAGACAGACACATCTTACTGAAG 314
Qy 21 AsnTyrIleuLysCysArgThrLysLysSerValGlnGluLeuLeuPheProLeu 40
Db 315 AATTACTTAATTAATGCAGAACCAAAAGAGTAGTGTTCAGGAAATTTCTTTTCCACTA 374
Qy 41 PhePheLeuPheTyrLeuIleLeuLeuSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 375 TTTTATTTATTTGGTTTAATTAATTAATGATGATGATGATGATGATGATGATGATGATG 434
Qy 61 ValProAsnLysGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuLys 80
Db 435 GTGCTAATATAGAACTCAATCTCTATGCAAGTTTACTCTTCTAATCTAATCTTCTTGA 494
Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAsnHisLeu 100
Db 495 TATACTCCAGTGACTAATATTAAGCAGCATCATGCAAGAAAGTGTCTCTGATCATCTA 554
Qy 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
Db 555 CCTGATGCTAATATTACTGAAGAATATACAAATGAAAGAAATGTTAATCATCCAGTCTC 614
Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
Db 615 TCTAAGCCGAGCAACTTTGAGTGTGTGTTTCAAGAGACTCCATGCTCTTAAGAACTCGT 674
Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
Db 675 TTTTCTCTGATGATGATCCAGTATCTTCTATTTATATGATGATTCAGAGCTGGCTGTCA 734
Qy 161 LysSerCysGluAlaAlaGlnTyrTyrSerSerGlyPheThrValLeuGlnAlaSerIle 180
Db 735 AAATCATGTGAGGCTGCTCAGTACTGTCTCCTCAGGTTTCAAGTTTACAAAGCATCCATA 794
Qy 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTyrLysGluLeuGluSerThr 200
Db 795 GATGCTGCCATTATACAGTTGAAGCAACATGTTTCTCTTTGGAGAGAGCTGGAGTCACT 854
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
Db 855 AAAGCTGTTATATGGGAGAACTGTGTTGTAGAAATAGATACCTTTTCCCGAGGAGTA 914
Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
Db 915 ATTTTATATACCTTAGTTATAGCATTTTCACTTTTGGATCTTTTGGCAATTCATATC 974
Qy 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
Db 975 GTAGCAGAAAAAGAAAAAATAAGAAATTTTAAAGATAATGGACTTCATGATACT 1034
Qy 261 AlaPheTyrLeuSerTyrValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeu 280
Db 1035 GCCTTTTGGCTTCTCTCTATATACAAAGTTTAAATTTTCTTATGTCCTTCTT 1094
```



QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
 DB 1921 CCAACAGCTGGGAATGGACCCCTGTTCTCGACATATGTATGATATCTTTTAAATAACAGA 1980  
 QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 DB 1981 AAAGCCAATCGGCTGACAGTGTTCAGTACTCATTTTCATGGATGAGCTGACATTCCTGCA 2040  
 QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
 DB 2041 GATAGAAAGCTGTGTATATCACAGAAATGCTGAATGTGTGTTCTTCATGTCTCTC 2100  
 QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
 DB 2101 AAAGTAAATGGGGATCGCTACCGCTGAGCATGTACATAGACAAATATGTGCCACA 2160  
 QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
 DB 2161 GAATCTCTTCTTCATGCTTTAAACACATATACCTGGAGCTCTTTTATACACAGAT 2220  
 QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
 DB 2221 GACCAACAATGTTGTATAGCTTGCTTTCAAGGACATGGACAAATTTTCAGGTTTGT 2280  
 QY 761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
 DB 2281 TCTGCCCTAGACAGTCATCAATTTGGGTGGCATTTCTTATGGGGTTTCCATGACGACT 2340  
 QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
 DB 2341 TTGGAGAGAGTATTTTAAAGCTAGAGTTGAAGCAGAAATGACCAAGCAGATATATAGT 2400  
 QY 801 ValPheThrGlnGlnProLeuGlnGluMetAspSerLysSerPheAspGluMetGlu 820  
 DB 2401 GTATTTTACTCAGCAGCCACTGGAGGAAGAAATGGATTCAAAATCTTTTGTGAATGGAA 2460  
 QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
 DB 2461 CAGAGCTTACTTATCTTCTGAAACCAAGGCTKCTCTAGTCAGCACCATGAGCCTTTGG 2520  
 QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
 DB 2521 AAACAACAGATGATACAAATAGCAAGTTTCAATTTCTTTACCTTGAAACGCTGAAAGTAA 2580  
 QY 861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
 DB 2581 TCAGTGAGATCAGTGTTGCTTCGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2640  
 QY 881 LeuValHisSerPheLysAsnAlaValProIleLysLeuValProAspLeuTyr 900  
 DB 2641 TTGGTTTCATCAGTCTTTTAAATATGCTGTGTTCCCATCAAACTTGTTCAGACTATAT 2700  
 QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
 DB 2701 TTTCTAAACCTCGAGACAAACACATAAATAAACAAGTCTGCTTCTTCAAAATCTT 2760  
 QY 921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
 DB 2761 GCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACAAGCCAGAACATATGTTGACG 2820  
 QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
 DB 2821 ATGATTAATGACAGTACATGATGATCGTGGCTCCCAATAGTGGGCTTTAAATGTRTG 2880  
 QY 961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
 DB 2881 CATTCGAAAGGACTATGTTTGTGAGCTGTTTTCACAGTACTATGTTTATCTTTA 2940  
 QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
 DB 2941 CCTATATTAGTGAATCATTTAGTAATACTATCTTTTATCATTTAAATGTGACTGAAC 3000

QY 1001 IleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLysIleGlu 1020  
 DB 3001 ATCCAGATCTGGATACCCCATTTCTTCAAGAAATTTACTGATATAGTTTAAAAATTGAG 3060  
 QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProProTyrPheAla 1040  
 DB 3061 CTGTATTTTCAAGCAGCTTTGCTTGGAAATCAITGTTACTGCAATGCCACCTTACTTTGCC 3120  
 QY 1041 MetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060  
 DB 3121 ATGGAAATGCGAGAAATCAATAGACTCAAGCTTAYACTCAACTTAAACTTTTCAGGTCT 3180  
 QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePheIleIle 1080  
 DB 3181 TTCCCATCTGCATATTTGGATTGGACAGCTGTTGTGATATCCCTTATTTTATCAT 3240  
 QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
 DB 3241 CTATTTTGTATGCTAGCAAGCTTATTTGGCATTTTCATATGAGTATATTTTATCTGTA 3300  
 QY 1101 LysPheLeuAlaValValPheCysLeuLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
 DB 3301 AAGTTCCTTGTCTGTGTTTTCCTTATTTGGTTATGTTCCATCAGTTATTTCTGTTCACT 3360  
 QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSerPheIle 1140  
 DB 3361 TATATTGCTTCTTTTCACTTTTAAAGAAATTTTAAATACCAAGAAATTTTGGTCAITTATC 3420  
 QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
 DB 3421 TATTCGTGTCAGCTGTTGCTTGTATGCAATCCTGCAATCACTTCTTTTATGGATAC 3480  
 QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleProIleTyrProLeuLeu 1180  
 DB 3481 ACATTTGCACTATTTCTTCAATATGCTTTTGTATCATCATTCATCCATCTATCCACTCTA 3540  
 QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThr 1200  
 DB 3541 GGTTCCTGATTTCTTTCATAAAGATTTCTTGAAGAATGTACGAAAAATTTGGACACC 3600  
 QY 1201 TyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
 DB 3601 TATATTCATGGATAGCTTTTCAAGTATGTTTATATCGCTTACCTGACGTGTGACTG 3660  
 QY 1221 TrpIlePheLeuGlnTyrTyrGluLysLysTyrGlyArgSerIleArgLysAsp 1240  
 DB 3661 TGGATTTTCTCTTACAAATCTATGAGAAAAATATGAGGCGAGATCAATAAGANAAGAT 3720  
 QY 1241 ProphePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAsp 1260  
 DB 3721 CCCTTTTTCAGAAACCTTTTCAAGAGTCTTAAATAAGAGGTTCCAGNACCACAGAC 3780  
 QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
 DB 3781 AATGAGCATGAGATGAGATGTCATAAGCTGAAAGACTAAAGCTCAAGAGCTGATGGGT 3840  
 QY 1281 CysGlnCysGlnGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAsp 1300  
 DB 3841 TGGCAGTGTGTGTGAGGAGAAACCATCCATATGTCAGCAATTTGCATAAAGAAATATGAT 3900  
 QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSer 1320  
 DB 3901 GACAAGAAAGATTTTCTTCTTCAAGAAAGTAAAGAAAGTGGCACTAAATATCATCTCT 3960  
 QY 1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
 DB 3961 TTTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGTCCAAATGTTGCTGCGCAAAAGC 4020  
 QY 1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
 DB 4021 ACAATTAATATATTCCTGTTGTGTGATATTTGAACCACTTTCAAGCCAGGATTTTATGGA 4080  
 QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380

```
; PRIOR FILING DATE: 2000-10-10
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4929
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-971-121-1

Alignment Scores:
Pred. No.: 0 Length: 4929
Score: 8395.00 Matches: 1637
Percent Similarity: 99.70% Conservative: 0
Best Local Similarity: 99.70% Mismatches: 5
Query Match: 99.63% Indels: 0
DB: 9 Gaps: 0

US-10-090-458-5 (1-1642) x US-09-971-121-1 (1-4929)

QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20
DB 1 ATGTCATGCAATAGGAGGTAGAGTTTGGAGACAGACAGACACATCTCTACTGAAG 60
QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerSerValGlnGlnIleLeuPheProLeu 40
DB 61 AATTACTTAATTAATGCGAAGCAAAAGAGTAGTGTTCAGGAATTCCTTTTCCACTA 120
QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60
DB 121 TTTTTTTTATTTGGTTAATTAATAGCATGATGATCCAAATAAGAAATATGAAGAA 180
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80
DB 181 GTGCCATATATAGAACTCAATCCTATGACCAAGTTACTCTTTCTAATCTAATCTTTGGA 240
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
DB 241 TATACTCCAGTGACTAATATTAACAGCAGCATCATGCAAGAAAGTGTCTACTGATCATCTA 300
QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
DB 301 CCTGATGTCATTAATCTGAGAGATATACAAATGAAAGAAATGTTAACTCCAGTCTC 360
QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
DB 361 TCTAAGCCGAGCAACTTTGTAGGTGTGGTTTCAAGACTCCATGCTCCTATGAATTCGT 420
QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
DB 421 TTTTTCCTGATATGATTCAGATATCTTCTATTTATATATGATTCAGAGCTGGCTGTTC 480
QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180
DB 481 AAATCATGTGAGGCTGCTCAGTACTGTGTCCTCAGCTTTTCAAGTTTACAGCATCCATA 540
QY 181 AspAlaAlaIleIleGluLeuLysThrAsnValSerIleTyrLysGluLeuGluSerThr 200
DB 541 GATGCTGCCATATACAGTTGAAGACCAATGTTCTCTTTGAGAGGAGCTGGAGTCAACT 600
QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
DB 601 AAAGCTGTTATATGGGAGAAACTGCTGTGTAGAAATAGATACCTTTCCCCGAGGAGTA 660
QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
DB 661 ATTTTAATATACCTAGTTATAGCATTTTTCACCTTTTGGATATCTTTTGGCAATTCATATC 720
QY 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
DB 721 GTAGCAGAAAGAAAGAAAGAAATTAAGAAATTTTAAGATAATGGGACTTCATGACTACT 780
QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280
DB 841 ATGGCAGTCATTTGGCAGACAGCTTCTTTGTTATTTCTCAAAAGTAGCAGCATTTGTGATATTT 900
QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320
DB 901 CTGCTTTTCTTCTTATGATATCATCTGATTTATTTTGTCTTATGCTGACACTCTT 960
QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340
DB 961 TTTTAAAAATCAAAACATGTGGGAATAGTTGAATTTTGTGTACTGTGGCTTTGGATTT 1020
QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360
DB 1021 ATTGGCTTTATGATATCTCATAGAAAGTTTTCCAAATCGTTAGTGTGGCTTTTCAGT 1080
QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380
DB 1081 CCTTCTGTCACTGTACTTTTGTGATTTGTTGATTTGCACAGTCTATGCAATTTAGAGATTTT 1140
QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400
DB 1141 AATGAAGGTGCTTCATTTTCAAAATTTGACTGCGAGGCCCATATCTCTAATTTACAAT 1200
QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420
DB 1201 ATCATGCTCACACTTAAATAGTATATTTCTATGCTCTTGGCTGTCTATCTTGATCAAGTC 1260
QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440
DB 1261 ATTCCAGGGGAATTTGGCTTTACGGAGATCATCTTTATATTTTCTGAAGCTTCATATTGG 1320
QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnClyAsnIleSer 460
DB 1321 TCAAGAGCAAAAGAAATTTATGAGGAGTTATCAGAGGGCAATGTTAATCGAAATATTAGT 1380
QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480
DB 1381 TTTAGTGAATTTATGAGCCAGTTTCTTCAGAAATTTGAGGAAAGAGCCATAGAAT 1440
QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500
DB 1441 AGTGGTATTTCAAGACATACAGAAAGAGGGTGAATTTGTGGAGGCTTTGAGAAATTTG 1500
QY 501 SerPheAspIleTyrGluGluGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520
DB 1501 TCATTTGACATATATGAGGCTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGGAAG 1560
QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540
DB 1561 AGTACATTTGATGAATATCTTTGTGACTCTGCCACCTTCTGATGGGTTTGCATCTATA 1620
QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560
DB 1621 TATGGACAGAGCTCTCAGAAATAGATGAATTTGTTGAAGCAAGAAATAGTTGGCAT 1680
QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580
DB 1681 TGTCACAGTATAGATATACATTTGATGTTTTGACAGTAGAAGAAATTTATCAATTTTG 1740
QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600
DB 1741 GCTTCAATCAAGGGATACCAGCCCAATATAATACAGAAAGTGCAGAGGTTTACTATA 1800
QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620
DB 1801 GATTTAGACATGCAGCATATCAAGATAACCAAGCTAAAAATTAAGTGGTGTCAAAA 1860
QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuAspGlu 640
DB 1861 AGAAAGCTGTCTATTAGGAATTCGTCTTGTGGAAACCCCAAGATATCTGCTGTAGTAA 1920
```

1001 IleGlnIleThrSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
Db ATCCAGATCTGGAGTACCCCACTCTTCAAGAAATTAAGTATAGTATTAAGATGAG 4070  
1021 LeuTyPheGlnAlaAlaLeuLeuGlyIleValThrAlaMetProTyPheAla 1040  
Db CTGTATTTTCAACAGCCTTGTGTGAATCAATGTTACTGCAATGCCACTTACTTTGCC 4130  
1041 MetGluAsnAlaGluAsnHisIleLysAlaTyThrGlnLeuLysLeuSerGlyLeu 1060  
Db ATCGAAATATGCAGAGATCATAGATCAAGCTTATATCACTTAAACTTTTCAGGTCTT 4190  
1061 LeuProSerAlaTyThrPheGlyGlnAlaValAlaValAspIleProLeuPhePheIle 1080  
Db TTGCCATCTGCATATTGGATTTGACCAAGCTGTTGTTGATATCCCTTATTTTATCAT 4250  
1081 LeuIleLeuMetLeuGlySerLeuAlaPheHisTyGlyLeuTyThrVal 1100  
Db CTATTTTGTATGCTAGAGAGCTTACCTGGCAATTCATATGGATATATTTTATCTGTA 4310  
1101 LysPheLeuAlaValValPheCysLeuIleGlyTyThrValProSerValIleLeuPheThr 1120  
Db AAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4370  
1121 TyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheThrSerPheIle 1140  
Db TATATTGCTTCTTCCCTTTAAGAAATTTTAAATACCAAGAAATTTGCTCATTTATC 4430  
1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTy 1160  
Db TATTCGTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4490  
1161 ThrIleAlaThrIleLeuHisTyThrAlaPheCysIleIlePheProLeuLeu 1180  
Db ACAATTGCAACTATTCTTCAATGCTTTTGTATCATCATTCATTCATTCATTCATTC 4550  
1181 GlyCysLeuLeuSerPheIleLysIleSerTyPheAsnValArgLysAsnValAspThr 1200  
Db GGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4610  
1201 TyrAsnProThrAspArgLeuSerValAlaValIleSerProTyLeuGlnCysValLeu 1220  
Db TATAATCCATGGGATAGGCTTTTCAAGAGTCTTAAATATAGAGCTTCCAGAACCCAGAC 4790  
1221 TrpIlePheLeuLeuGlnTyTyGlyLysTyGlyGlyArgSerIleArgLysAsp 1240  
Db TGCATTTTCTCTTACAATACTATGAGAAAAATATGAGAGCAGATCAATAAGAAAGAT 4730  
1241 ProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProAsp 1260  
Db CCTTTTTCAGAAACCTTTTCAAGAGTCTTAAATATAGAGCTTCCAGAACCCAGAC 4790  
1261 AsnGluAspGluAspGluValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
Db AATGAGGATGAAGATGAAGATCTCAAGCTGAAAGCTAAAGGTCAAGAGCTGATGGGT 4850  
1281 CysGlnCysGluGlnLysProSerIleMetValSerAsnLeuHisLysGluTyAsp 1300  
Db TGCAGTGTGTGAGGAGAAACCATCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4910  
1301 AspLysLysAspPheLeuLeuSerArgLysValLysValLysValAlaThrLysTyThrIleSer 1320  
Db GACAAGAAAGATTTTCTTTTCAAGAAAGTAAAGAAAGTGGCACTAAATATCATCTCT 4970  
1321 PheCysValLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSer 1340  
Db TTCTGTGTGAAAAAGAGAGATCTTAGGACTATTGGGTCCAAATGGTCTGGCAAAAGC 5030  
1341 ThrIleIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
Db ACANTTATTATTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5090

1361 AspTySerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyTyCysPro 1380  
Db GATTATTCTTCAGAGACAGAGTGAAGATGATGATTCACTGAAGTGTATGGTACTGCTCT 5150  
1381 GlnIleAsnProLeuTyProAspThrThrLeuGlnGluHisPheGluIleTyGlyAla 1400  
Db CAGATAAACCTTGTGTGGCCAGATACATACATTCAGAGAAATTTTGAATTTATGAGCT 5210  
1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420  
Db GTCAAGGAATGAGTCAAGTGCATGAAAGAGTCAATAGTCGAATACACATGCACCTT 5270  
1421 AspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyLysArgLys 1440  
Db GATTAAAGAAACATCTTCAGAGACTGTAAAGAACTACCTGCAGAAATCAACACGAAAG 5330  
1441 LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460  
Db TTGTGTTTCTCTAAGTATGCTAGGAACTCTCAGATTAATTTGCTAGATGAACCATCT 5390  
1461 ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480  
Db ACAGGTATGATCCCAAGCCAAACAGCAGCATGTGGCGCAATTCGAACCTGCATTTAAA 5450  
1481 AsnArgLysArgAlaAlaIleLeuThrThrHisTyMetGluGluAlaGluAlaValCys 1500  
Db AACAGAAAGCGGCTGCTATTCGACCATCTATATGAGGAGGAGGAGGCTGCTGT 5510  
1501 AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520  
Db GATCGATAGCTATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5570  
1521 LysSerLysPheGlyLysGlyTyPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540  
Db AAGATTAATTTGGAAGAGCTACTTTTGGAAATTAATTTGAAGNCTGATAGAAAC 5630  
1541 LeuGluValAspArgLeuGlnArgGluIleGlnTyThrPheProAsnAlaSerArgGln 1560  
Db CTAGAGTAGACCCCTTCAAGAGAAATTCAGTATATTTTCCCAATGCAAGCCGCTCAG 5690  
1561 GluSerPheSerIleLeuAlaTyLysIleProLysGluAspValGlnSerLeuSer 1580  
Db GAAAGTTTCTCTCTATTTTGGCTTATAAATTCCTAAGAAAGATGTTCACTCCCTTCA 5750  
1581 GlnSerPhePheLysLeuGluAlaLysHisAlaPheAlaIleGluGluTySerPhe 1600  
Db CAATCTTTTAAAGCTGGAAGAGCTAAACATGCTTTTGCCATTCGAAGATATAGCTTT 5910  
1601 SerGlnAlaThrLeuGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620  
Db TCTCAAGCAACATTTGGAACAGGCTTTTGTAGAACTCATAAAGAACCAAGAGGAGAT 5970  
1621 AsnSerCysGlyThrLeuAsnSerThrLeuTyTrpGluArgThrGlnGluAspArgVal 1640  
Db AATAGTTGTGAACCTTTAAACAGCACACTTTTGTGGGAGCAACACAAAGAGATAGATA 5930  
1641 ValPhe 1642  
Db GTATTT 5936

## RESULT 3

US-09-971-121-1

; Sequence 1, Application US/09971121

; Patent No. US20020111477A1

; GENERAL INFORMATION:

; APPLICANT: Hu, Yi

; TITLE OF INVENTION: Nephomichy, Boris

; TITLE OF INVENTION: Same

; FILE OF INVENTION: LEX-0250-USA

; CURRENT APPLICATION NUMBER: US/09/971,121

; CURRENT FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: US 60/239,629

Human Transporter Proteins and Polynucleot

1791 GCCTTTGGCTTTCCTGGGTTCTTCTATATACAAAGTTTAAATTTCTTATGTCCTCTCT 1850  
181 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
1851 ATGGCAGTCAATGCGACAGCTCTTTGTTATTTATTCCTCAAGTAGCAGCAATTCGTATATT 1910  
301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
1911 CTGCTTTTTCCTTATGGAATATCATCTGTATTTTTCCTTAAATGCTGACACTCTT 1970  
321 PheLysGlySerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
1971 TTTAAAAATCAAAACATGTGGGAATAGTTGAAATTTTTTGTACTGTGGCTTTTGGATTT 2030  
341 IleGlyLeuMetIleIleLeuIleGlySerPheProLysSerLeuValThrLeuPheSer 360  
2031 ATGGGCTTATGTAATCTCATAGAAAGTTTCCCAATCTGTAGTGGCTTTTTCAGT 2090  
361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPhe 380  
2091 CCTTTCTGTCACTGTACTTTTCAAAATTTGACTGAGGCCCATATCTCTTAATATTACAATT 2150  
381 AnGluGlyAlaSerPheSerHisLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
2151 AATGAAGTGGCTTCATTTTCAAAATTTGACTGAGGCCCATATCTCTTAATATTACAATT 2210  
401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValThrLeuAspGlnVal 420  
2211 ATCATGCTCACACTTAATAGTATATTTCTATGTCCTCTTGGCTGTCTATCTTGATCAAGTC 2270  
421 IleProGlyGluPheGlyLeuArgArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
2271 ATTCCAGGGGAATTTGGCTTACGGAGATCATCTTTATATTTTCTGAAGCCTTCATATTGG 2330  
441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAspGlyAsnIleSer 460  
2331 TCAAGAGTAAAGAAATATGAGGAGTTATCAGAGGCCAATGTTAAATGGAATATTAGT 2390  
461 PheSerGluIleGluProValSerSerGluPheValGlyLysGluAlaIleArgIle 480  
2391 TTTAGTGAATATTATGAGCCAGTTCTTTCAGAAATTTGTAGGAAAAAGAGCCATAGAAAT 2450  
481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
2451 AGTGGTATTCAAGAGCATACAGAAAGAGGGTGAATATGGAGGCTTTGAGAAATTTG 2510  
501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
2511 TCATTTGACATATATGAGGCTCAGATTACTGCTTACTTGGCCACAGTGGAAACAGGAAAG 2570  
521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSerIle 540  
2571 AGTACATTGATGAATTTCTTTGGAGCTCTCCCCACCTTCTGATGGGTTTGCATCTATA 2630  
541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
2631 TATGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATATGATTGGCAATT 2690  
561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
2691 GTCCACAGTTAGATATACACTTTTGTATTTGACAGTAGAAGAAATTTATCAATTTTG 2750  
581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
2751 GCTTCAATCAAGGGATACCCCAACCAATATAATACAAAGAGTGCAGAGGTTTACTA 2810  
601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
2811 GATTTAGACATCCAGACTATCAAGATTAACCAAGCTAAAAAATTAAGTGGTGGTCAAAA 2870  
621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640

2871 AGAAAGCTGTCAITTAGGAATTTGCTTTCTTGGGAACCCAAAGATACCTGCTGTAGTGAA 2930  
641 ProThrIlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
2931 CCAACAGCTGGAATGGACCCCTGTCTCGACATATTGTATGGAATCTTTTAAAAATCAGA 2990  
661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
2991 AAAGCCAAATCGGTGACAGTGTTCAGTACTCATTTCAATGATGAAGTGACATCTTTGCA 3050  
681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMetPheLeu 700  
3051 GATAGGAAGCTGTGATATCACAAAGAAATGCTGAAATGTGTGTTCTTCAATGTTCTCTC 3110  
701 LysSerLysTyrGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
3111 AAAGTAAATGGGGGATCGGCTACCGCTGAGGATGTACATAGACAAATATTGTGCCACA 3170  
721 GluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnAsn 740  
3171 GAATCTCTTCTTCACTGTGTTAAACAACATATACCTGGAGCTACTTTATTACAACAAAT 3230  
741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
3231 GACCAACAATTTGTGTATAGCTTGGCTTTCAAGGACATGGACAAATTTTCAGGTTTGT 3290  
761 SerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
3291 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCTATTCTTATGTTGTTTCCATGACGACT 3350  
781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
3351 TTGGAAGACGTATTTTAAAGCTAGAGTTGAAGTGAAGCAAAATTCACCAAGCAGATATAGT 3410  
801 ValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGluMetGlu 820  
3411 GTATTTACTCAGCAGCCACTGGAGGAGAAATCGATTCAAAATCTTTTGATGAATGGAA 3470  
821 GluSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTrp 840  
3471 CAGAGCTTACTTATCTTCTGAAACCAAGGCTTCTCTAGTAGAGCACCATGAGCCTTTGG 3530  
841 LysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGluSerLys 860  
3531 AAACACAGATGATACATAGCAAGTTTCATTTCTTACCTTGAACAGTGAAGTAA 3590  
861 SerValArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
3591 TCAGTGAGATCAGTGTGCTTCTGCTTTTAAATTTTTTTTTCACAGTTTATATGTTT 3650  
881 LeuValHisHisSerPheLysAsnAlaValProIleLysLeuValProAspLeuTyr 900  
3651 TTGGTTCACTCAGCTTTTAAATATGCTGGTTCCCATCAAACTTGTCTCCAGACTTATAT 3710  
901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSer 920  
3711 TTTCTAAACCTGGAGCAACCAACATATAACAAACAGTCTGCTTCTTCAAAATCTT 3770  
921 AlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
3771 GCTGACTCAGATATCAGTATCTTATAGCTTTTTCACAAAGCCAGCAACATAATGAGG 3830  
941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsnValMet 960  
3831 ATGATTAATGACAGTACTATGATCCGTGGCTCCCATAGTGGGCTTTAAATGTGATG 3890  
961 HisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyrSerLeu 980  
3891 CATTCAAAAAGGACTATGTTTTCAGCTGTTTTCACAGTACTATGTTTATCTTTA 3950  
981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThr 1000  
3951 CCTATTATAGTGAATATCAATAGTAATACTACTATCTTTATCATTTAAATGTGACTGAACC 4010

```

Db 4314 GTCAAGGAATGAGTCAAGTGACATGAAAGAGTCAATAGTCGAATACACATGCACTT 4373
Qy AspLeuLysGluHisLeuGlnLysThrValLysLeuProAlaGlyLeuLysArgLys 1440
Db 4374 GATTTAAAGAACATCTTCAGAGAGCTGTAAGAAACTACCTGCGAGGATCAACAGAAAG 4433
Qy LeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSer 1460
Db 4434 TTGTGTTTCTCTAAGTATGCTAGGGAATCTCAGATTACTTCTCAGATTACTTGTAGATGAACCATCT 4493
Qy ThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLys 1480
Db 4494 ACAGGTATGATCCCAAGGCCAAACAGCAGCATGTGGCGCAATTCGAATGCATTTAAA 4553
Qy AsnArgLysArgAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCys 1500
Db 4554 AACAGAAAGCGGCTCTATTCTGACCACTCACTATATGAGGAGGCGAGAGCTGTCGT 4613
Qy AspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeu 1520
Db 4614 GATCGAGTAGCTATCATGTGCTGGCGAGTTAAGATGTATCGGAACAGATCAACATCTTA 4673
Qy LysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsn 1540
Db 4674 AAGAGTAATTTGGAAAGGCTACTTTTGGAAATTAATTTGAAGACTGGATAGAAAC 4733
Qy LeuGluValAspArgLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGln 1560
Db 4734 CTAGAAGTAGACGCGCTTCAAGAGAAATTCAGTATATTTCCCAATCAAGCGGTGAG 4793
Qy GluSerPheSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSer 1580
Db 4794 GAAAGTTTTCTCTATTGCTTTGGCTTATATAATTTCTTAAGAGATGTTTCAGTCCCTTCA 4853
Qy GlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPhe 1600
Db 4854 CAATCTTTTTTAAGCTGGAAGAGCTAAACATGCTTTTCCCATTAAGAGATGTTTCAGTCCCTTCA 4913
Qy SerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAsp 1620
Db 4914 TCTCAAGCAACATGGAGACAGGTTTTGTAGAACTCCTTAAGAACAGAGGAGAGAT 4973
Qy AsnSerCysGlyThrLeuAsnSerThrLeuTrpTrpGluArgThrGlnGluAspArgVal 1640
Db 4974 AATAGTTGTGGAATTTAAACAGCACACTTTGGTGGGAACGAACACAGAGATAGATGA 5033
Qy ValPhe 1642
Db 5034 GTATTT 5039

RESULT 2
US-10-005-338B-1
; Sequence 1, Application US/10005338B
; Publication No. US20030044895A1
; GENERAL INFORMATION:
; APPLICANT: DENEFE, Patrice
; APPLICANT: ROSIER-MONTUS, Marie-Francoise
; APPLICANT: PRADES, Catherine
; APPLICANT: ARNOULD-REGUINE, Isabelle
; APPLICANT: DUVERGER, Nicolas
; APPLICANT: ALLIKMETS, Rando
; APPLICANT: DEAN, Michael
; TITLE OF INVENTION: NUCLEIC ACIDS OF THE HUMAN ABCA5, ABCA6, ABCA9, AND ABCA10 GENES
; FILE REFERENCE: ABCA5, 6, 9, 10
; CURRENT APPLICATION NUMBER: US/10/005,338B
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/263,231
; PRIOR FILING DATE: 2001-01-23
; PRIOR APPLICATION NUMBER: FR 00403440.1
; PRIOR FILING DATE: 2000-12-07
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.1

```

```

; SEQ ID NO 1
; LENGTH: 6525
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 4449
; OTHER INFORMATION: n=unknown, may be a g o r c o r t
US-10-005-338B-1

Alignment Scores:
Pred. No.: 0 Length: 6525
Score: 8422.00 Matches: 1641
Percent Similarity: 99.94% Conservatives: 0
Best Local Similarity: 99.94% Mismatches: 1
Query Match: 99.95% Indels: 0
DB: 14 Gaps: 0

US-10-090-458-5 (1-1642) x US-10-005-338B-1 (1-6525)
Qy 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20
Db 1011 ATGTCCACTGCAATTAGGGAGGTAGGAGTTTGAGACAGACCAGAACACTTCTACTGAAG 1070
Qy 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGluIleLeuPheProLeu 40
Db 1071 AATTACTTAAATTAATGACAGAACCAAAAGAGTAGTGTTCAGGAATTTCTTTTCCACTA 1130
Qy 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 1131 TTTTCTTATTTTGGTTTAAATTAATTAATGATGATGATGATGATGATGATGATGATGATG 1190
Qy 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80
Db 1191 GTGCTAATTAATGAACTCAATCTCTATGGACAGTTTACTCTTCTTAATCTAATCTTGG 1250
Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
Db 1251 TATATCTCCAGTGAATAATTAATCAAGCAGCATCATGACAGAAAGTGTCTACTGATCATCTA 1310
Qy 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120
Db 1311 CCTGATGCTAATAATCTGAGGAATATACAAATGAAAGAAATGTTTAACATCAGTCTC 1370
Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140
Db 1371 TCTAAGCCGACACTTTGTAGGTGTGTTTTCAAAGACTCCATGCTCTATGAATTCGT 1430
Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160
Db 1431 TTTTCTCTGATATGATTCAGTATCTTCTATTATATGATGATCAAGAGCTGCTGTCTCA 1490
Qy 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180
Db 1491 AATCATGTGGGCTGCTCAGTACTGCTCCTCAGTGTTCACAGTTTTCACAGCATCCATA 1550
Qy 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200
Db 1551 GATGCTGCCATTATACAGTTGAAGACCAATGTTTCTCTTGGAGAGAGCTGGAGTCAACT 1610
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220
Db 1611 AAACCTGTTATATGGAGAAACCTGCTGTGTGAGAAATAGATACCTTTCCCGAGAGGTA 1670
Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240
Db 1671 ATTTTAAATATACCTAGTATTAGCAATTTTCCACTTTTGGTACTTTTGGCAATTCATATC 1730
Qy 241 ValAlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThr 260
Db 1731 GTAGCAGAAAGAAAGAAAGAAATTTTAAAGATATAGGAGACTTCATGATACT 1790
Qy 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280

```

QY 681 AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerMetPheLeu 700  
DB 2154 GATAGGAAGCTGTGATATCATCAAGGAATGCTGAAATGTGTGGTTCTTCAATGTTCCTC 2213  
QY 701 LysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThr 720  
DB 2214 AAAAGTAAATGGGGGATCGGCTACCGCTGAGCATGTACATAGACAAATATTGTGCCACA 2273  
QY 721 GluSerLeuSerSerLeuValLysGlnHisIleProGlnValaThrLeuLeuGlnGlnAan 740  
DB 2274 GAATCTCTTCTTCACTGGTTAAACAACATATACCTGGAGCTACTTTATTACACAGAAT 2333  
QY 741 AspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGlyLeuPhe 760  
DB 2334 GACCAACAACCTGTGTATAGCTTGCCTTCAAGGACATGGACAAATTTTCAGGTTTGTT 2393  
QY 761 SerAlaLeuAspSerHisSerAanLeuGlyValIleSerTyrGlyValSerMetThrThr 780  
DB 2394 TCTGCCCTAGACAGTCATTCAAATTTGGGTGTCAATTTCTATGGTGTTCCTCAAGCAGCT 2453  
QY 781 LeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSer 800  
DB 2454 TTGGAAGACGTAATTTTAAAGCTAGAGTTGAAGTCAGAGAAATTCACCAACAGATATAGT 2513  
QY 801 ValPheThrGlnGlnProLeuGlnGluGluMetAspSerLysSerPheAspGluMetGlu 820  
DB 2514 GTATTTACTCAGCAGCCACTGGAGGAGAAATGGATTCAAAATCTTTTGTGAAATGGAA 2573  
QY 821 GlnSerLeuLeuIleLeuSerGluThrLysAlaSerLeuValSerThrMetSerLeuTyr 840  
DB 2574 CAGAGCTTACTTATCTTCTGAAACCAAGGCTTCTCTAGTGAGCACCATGAGCCTTGG 2633  
QY 841 LysGlnGlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLys 860  
DB 2634 AAAACAACAGATGATACAAATAGCAAGTTTCATTTCTTTACCTTGAAACGCGAAAGTAA 2693  
QY 861 SerValArgSerValLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPhe 880  
DB 2694 TCAGTGAGATCAGTGTTCTTCTGCTTTTAAATTTTTCACAGTTCAGATTTTATGTTT 2753  
QY 881 LeuValHisSerPheLysAsnAlaValValProIleLysLeuValProAspLeuTyr 900  
DB 2754 TTGGTTTCATCAGCTCTTTTAAATATGCTGTGTTCCCATCAAACTTGTTCAGACTATAT 2813  
QY 901 PheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAanSer 920  
DB 2814 TTTCTAAACCTGGAGACAAACACATAAATACAAACCAAGTCTGCTCTTCAAAATCT 2873  
QY 921 AlaAspSerAspLysSerAspLeuIleSerPhePheThrSerGlnAsnIleMetValThr 940  
DB 2874 GCTGACTCAGATATCAGTGATCTTATAGCTTTTTCACCAAGCCAGAACATATGTGAGC 2933  
QY 941 MetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAanValMet 960  
DB 2934 ATGATTAATGACAGTACATGTATGTCGCTCCCATAGTGCAGCTTTTAAATGTGATG 2993  
QY 961 HisSerGluLysAspTyrValPheAlaValPheAsnSerThrMetValTyrSerLeu 980  
DB 2994 CATTGAAANAAGACTATGTTTGTGAGCTGTTTCAACAGTACTATGTTTATCTTTA 3053  
QY 981 ProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAanValThrGluThr 1000  
DB 3054 CCTATATTAGTGAATATCATTTAGTAATACTACTATCTTTATCAATTAATGTGACTGAAAC 3113  
QY 1001 IleGlnIleTrpSerThrProPheGlnGluIleThrAspIleValPheLysIleGlu 1020  
DB 3114 ATCCAGATCTGGAGTACCCTATCTTTCAAGAAATTTACTGATATAGTTTAAATGAG 3173  
QY 1021 LeuTyrPheGlnAlaAlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAla 1040  
DB 3174 CTGTATTTTCAAGCAGCTTGTGTGAATCATTTGTACTGCAATGCCACCTTACTTTGCC 3233  
QY 1041 MetGluAanAlaGluAanHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeu 1060

DB 3234 ATGGAAATGCAGAGATCATAGATCAAGCTTATATCAACTTAACTTCAAGTCTT 3293  
QY 1061 LeuProSerAlaTyrTrpIleGlyGlnAlaValAspIleProLeuPhePheIleIle 1080  
DB 3294 TTGCCATCTGCATATTGGATGGCAAGCTGTGTGTGATATCCCTTATTTTATCATT 3353  
QY 1081 LeuIleLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrVal 1100  
DB 3354 CTTATTTTGTAGCTAGGAGCTTATGGCATTTTCATTTATGATATATTTTATCATGTA 3413  
QY 1101 LysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThr 1120  
DB 3414 AAGTCTCTTGTGTGGTTTTTGTCTTATTTGGTTATCTTCATCAGTATTTCTGTCACT 3473  
QY 1121 TyrIleAlaSerPheThrPheLysLysIleLeuAanThrLysGluPheTrpSerPheIle 1140  
DB 3474 TATATTGCTTCTTTCACCTTTAGAAAATTTTAAATACCAAGAAATTTTGGTCATTATC 3533  
QY 1141 TyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr 1160  
DB 3534 TATTCTGTGGCAGCGTTGGCTTGTATTGCAATCACTGAAATAACTTTCTTTATGGGATAC 3593  
QY 1161 ThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeu 1180  
DB 3594 ACAATTGCACTATTTCTTCAATATGCTTTTGTATCATCATTCATCCACTTCTCTA 3653  
QY 1181 GlyCysLeuIleSerPheIleLysIleSerTrpLysAanValArgLysAanValAspThr 1200  
DB 3654 GGTTCCTGATTTCTTTCAATAAGATTTCTTGAAGAATGTACGAAAAAATGTGCACACC 3713  
QY 1201 TyrAanProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeu 1220  
DB 3714 TATAATCCATGGGATAGCTTTCAGTAGCTGTATATCGCTTACCTGCGAGTGTACTG 3773  
QY 1221 TrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyGlyArgSerIleAglLysAsp 1240  
DB 3774 TGGATTTTCTCTTACATTAATGAGAAAAATATGGAGGCAGATCAATAGAAAAAGAT 3833  
QY 1241 ProPheArgAanLeuSerThrLysSerLysAanArgLysLeuProGluProAsp 1260  
DB 3834 CCCCCTTTTTCAGAAACCTTTCAACGAGCTTAAAAATAGGAAGCTTCCAGAACCCACAGAC 3893  
QY 1261 AsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGly 1280  
DB 3894 AATGAGATGAAGATGAAGATGTCAAAGCTGAAAGCTAAAGGTCAAGAGCTGATGGGT 3953  
QY 1281 CysGlnCysGluGluLysProSerIleMetValSerAanLeuHisLysGluTyrAsp 1300  
DB 3954 TGCCAGTGTGTGAGGAGAAACCATCCATTATGTCAGCAATTTGCATAAAGAATATGAT 4013  
QY 1301 AspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSer 1320  
DB 4014 GACAAGAAGATTTTCTTCTTCAAGAAAAATGAAAGAGTGGCAACTTAAATACATCTCT 4073  
QY 1321 PheCysValLysLysGlyGluIleLeuGlyLeuGlyProAanGlyAlaGlyLysSer 1340  
DB 4074 TTCTGTGTGAAAAAGAGAGAGATCTTAGGACTATTGGTCCAAATGGTGTGCGCAAGAC 4133  
QY 1341 ThrIleIleAanIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGly 1360  
DB 4134 ACAATTAATTAATTTCTGTTGGTGTATATGAACCAACTTCAGGCCAGGTATTTTAGGA 4193  
QY 1361 AspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysPro 1380  
DB 4194 GATTAATTTCTTCAGAGACAGAGATGATGATTCATCTGAAGTGTATGGGTACTGTCT 4253  
QY 1381 GlnIleAanProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAla 1400  
DB 4254 CAGATAAACCTTTGTGCCAGATACTATCATTTGAGAAACATTTTGAATTTATGAGCT 4313  
QY 1401 ValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeu 1420

## Alignment Scores:

Pred. No.: 0 Length: 5475  
 Score: 8426.00 Matches: 1642  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-10-090-458-5 (1-1642) x US-10-090-458-4 (1-5475)

QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLeuLys 20  
 Db 114 ATGTCCACATGCAATTAGGAGAGTAGGAGTTTGAGACAGACAGACACATCTACTGAAG 173  
 QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGluLeuLeuPheProLeu 40  
 Db 174 AATTACTTAATTAATTCAGAACCAAAAGAGTAGTGTTTCAGAAATCTTTTTCACATA 233  
 QY 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
 Db 234 TTTTCTTTTATTTGGTTAATATTAATATAGCATGATCATCCAAATAAGAAATATAGAA 293  
 QY 61 ValProAsnIleGluLeuLeuProMetAspLysPheThrLeuSerAsnLeuIleLeuGly 80  
 Db 294 GTCCCTTAATATAGAACTCAATCTATGGACAGTTTACTCTTCTTAATCTAATCTTGA 353  
 QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
 Db 354 TATACTCCAGTGACTAATATTAACAGCAGCATCATCGAAGAGTGCTACTCATCTA 413  
 QY 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
 Db 414 CCTGATGTCTAATATTAAGAAATATACAAATGAAAGAAAGAAATGTTAACTCCAGTCTC 473  
 QY 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
 Db 474 TCTAACCCGAGCACTTTGTAGTGTTGTTTCAPAGACTCCATGTCTATGAATTCGT 533  
 QY 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgIleGlyCysSer 160  
 Db 534 TTTTCTTCTGATATGATCCAGTATCTTCTATATATGATGATTCAGAGCTGGCTGTTCA 593  
 QY 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerGlyPheThrValLeuGlnAlaSerIle 180  
 Db 594 AAATCATGTGAGCTGCTCAGTACTGCTCTCAGGTTTACAGTTTACAGCATCCATA 653  
 QY 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
 Db 654 GATGCTGCCATTATACAGTTGAAGCAATGTTCTCTTTGGAAGAGCTGGAGTCAACT 713  
 QY 201 LysAlaValIleMetGlyGluThrAlaValValGluIleAspThrPheProArgGlyVal 220  
 Db 714 AAAGCTGTATATAGGAGAACTGCTGTTGTAGAAATAGATACCTTTCCCGAGGAGTA 773  
 QY 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
 Db 774 ATTTTAATATACCTAGTTATAGCATTTTACCTTTTGGATACCTTTTGGCAATTCATATC 833  
 QY 241 ValAlaGluLysGluLysIleLysGluPheLysIleMetClyLeuHisAspThr 260  
 Db 834 GTAGCAGAAAAAAGAAAAATAAAGAAATTTTAAAGATATGGAATTCATGATACT 893  
 QY 261 AlaPheTrpLeuSerTrpValLeuLeuTyrThrSerLeuIlePheLeuMetSerLeuLeu 280  
 Db 894 GCCTTTTGGCTTCTCTGGGTTCTTCTATATACAGTTTAAATTTTCTATGTCCTTCTT 953  
 QY 281 MetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleValIlePhe 300  
 Db 954 ATGGCAGTCATTCGACAGCTCTTTTGTATTTTCCCTCAAGTAGCAGCATTTGTATTT 1013  
 QY 301 LeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThrProLeu 320  
 Db 1014 CTGCTTTTCTTCTTATGGAATATCATCTGATATTTTGTCTTAAATGCTGACACTCTT 1073

QY 321 PheLysLysSerLysHisValGlyIleValGluPhePheValThrValAlaPheGlyPhe 340  
 Db 1074 TTTAAAAAATCAAAACATGTGGAAATAGTTGAATTTTGTACTGTGGCTTTTGGATTT 1133  
 QY 341 IleGlyLeuMetIleIleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSer 360  
 Db 1134 ATTGGCCTTATGATAATCTCATAGAAAGTTTCCCAATCGTTAGTGTGGCTTTTCAGT 1193  
 QY 361 ProPheCysHisCysThrPheValIleGlyIleAlaGlnValMethIleLeuGluAspPhe 380  
 Db 1194 CCTTTCTGTCACTGTACTTTTGTGATTGGTATTGCACAGTCATGCATTTAGAAATTTT 1253  
 QY 381 AsnGluGlyAlaSerPheSerAsnLeuThrAlaGlyProTyrProLeuIleIleThrIle 400  
 Db 1254 AATGAAGTCTCTCATTTTCAAATTTGACTCGAGGCCATATCTCTAA'TATTACAAT 1313  
 QY 401 IleMetLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyrLeuAspGlnVal 420  
 Db 1314 ATCATGTCTACATTAATATATCTATCTCTCTCTGCTCTCTATCTTGTATCAAGTC 1373  
 QY 421 IleProGlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLysProSerTyrTrp 440  
 Db 1374 ATTCACAGGGAAATTTGGCTTACGGAGATCATCTTTATATATTTCTGAAGCCTTCAATGG 1433  
 QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAlaAsnGlyAsnIleSer 460  
 Db 1434 TCAAAGAGCAAAAGAAATATAGAGAGTTATCAGAGGCAATGTTAATGGAATATTAGT 1493  
 QY 461 PheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaIleAspGile 480  
 Db 1494 TTTAGTGAATATTTAGCCAGTTTCTTCAGAAATTTGTAGGAAAGAGCCATTAAGAT 1553  
 QY 481 SerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeu 500  
 Db 1554 AGTGGTATTCAAGAGACATACAGAAAGAGGGTGAATGTGAGGCTTTGAGAAATTTG 1613  
 QY 501 SerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLys 520  
 Db 1614 TCATTTGACATATATAGGGTCAGATTACTGCTTACTTGGCCACAGTGGAAAGAGAAAG 1673  
 QY 521 SerThrLeuMetAsnIleLeuCysGlyLeuCysProSerAspGlyPheAlaSerIle 540  
 Db 1674 AGTACATTGATGAATATCTTTGTGACTCTGCCCACCTTCTGATGGTTTGCATCTATA 1733  
 QY 541 TyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGlyIle 560  
 Db 1734 TATGGACACAGAGTCTCAGAAATAGATGAATGTTTGAAGCAAGAAATGATTTGGCATT 1793  
 QY 561 CysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIleLeu 580  
 Db 1794 TGTCCACAGTTAGATATACACTTTGATGTTTTCACAGTAGAAGAAATTTATCAATTTG 1853  
 QY 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 Db 1854 GCITTCATCAAGGGATACCAGCCACATATATATCAAGAAAGTCAGAGAGTTTACTA 1913  
 QY 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 Db 1914 GATTTAGACATGACAGCTATCAAGATACCAAGCTATAAAAAATTAAGTGTGTGCAAAA 1973  
 QY 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
 Db 1974 AGAAAGCTGTCA'TTAGGAATTTGCTTCTTTGGGAACCCCAAGATACCTGCTGTAGATGA 2033  
 QY 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg 660  
 Db 2034 CCAACAGCTGAATGGAGCCCTGTTCTCGACATATTTGTATGGAATCTTTTAAATACAGA 2093  
 QY 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIleLeuAla 680  
 Db 2094 AAACCCAAATCGGGTCAGAGTGTTCAGTACTCAATTTTCATGATGCAAGCTGACATCTTGCA 2153



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 13, 2004, 00:30:55 ; Search time 860 seconds  
(without alignments)

7162.674 Million cell updates/sec

Title: US-10-090-458-5

Perfect score: 8426

Sequence: 1 MSTAIREYGVNRQRTLLK.....CGTLNSTLWERTQEDRVVF 1642

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2475585 seqs, 1875730760 residues

Total number of hits satisfying chosen parameters: 4951170

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10090458/runat 08042004 121554 25691/app query.fasta\_1.1799  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10090458 -CGN\_1\_309 @runat 08042004 121554 25691  
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LOGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

## RESULT 1

US-10-090-458-4  
; Sequence 4, Application US/10090458  
; Publication No. US20020123107A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Hongyun  
; APPLICANT: Kilinski, Ligia  
; APPLICANT: Le Bihan, Stephane  
; TITLE OF INVENTION: NOVEL ABCA5 TRANSPORTER AND USES THEREOF  
; FILE REFERENCE: 100103.403  
; CURRENT APPLICATION NUMBER: US/10/090,458  
; CURRENT FILING DATE: 2002-03-01  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 5475  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 25  
; OTHER INFORMATION: n = A,T,C or G  
US-10-090-458-4

## ALIGNMENTS

	1	8426	100.0	5475	13	US-10-090-458-4	
Sequence 4, Appli	1	8422	100.0	6525	14	US-10-005-338B-1	Sequence 1, Appli
Sequence 1, Appli	2	8395	99.6	5282	9	US-09-371-121-1	Sequence 5, Appli
Sequence 5, Appli	3	8395	99.6	5282	9	US-09-371-121-5	Sequence 3, Appli
Sequence 3, Appli	4	8309	98.6	4917	13	US-10-090-458-3	Sequence 3, Appli
Sequence 3, Appli	5	8309	98.6	4917	13	US-10-090-458-1	Sequence 3, Appli
Sequence 3, Appli	6	8112	96.3	4785	9	US-09-371-121-3	Sequence 3, Appli
Sequence 3, Appli	7	4930.5	58.5	3347	15	US-10-108-260A-160	Sequence 160, App
Sequence 160, App	8	4847.5	57.5	3268	15	US-10-094-749-984	Sequence 984, App
Sequence 984, App	9	4847.5	57.5	3268	15	US-10-005-338B-4	Sequence 4, Appli
Sequence 4, Appli	10	3234	38.4	6181	13	US-10-090-454-3	Sequence 3, Appli
Sequence 3, Appli	11	3200.5	38.0	4875	13	US-10-090-454-1	Sequence 1, Appli
Sequence 1, Appli	12	3200.5	38.0	5018	13	US-10-005-338B-3	Sequence 3, Appli
Sequence 3, Appli	13	3197.5	37.9	5981	14	US-10-005-338B-2	Sequence 2, Appli
Sequence 2, Appli	14	3124	37.1	5296	14	US-10-090-453A-1	Sequence 1, Appli
Sequence 1, Appli	15	3117	37.0	4854	13	US-10-090-453A-3	Sequence 3, Appli
Sequence 3, Appli	16	3117	37.0	5332	13	US-10-090-453A-1	Sequence 1, Appli
Sequence 1, Appli	17	3107.5	36.9	3928	10	US-09-822-846-98	Sequence 98, Appli
Sequence 98, Appli	18	2982	35.4	4798	15	US-10-093-463-137	Sequence 137, App
Sequence 137, App	19	2961.5	35.1	4766	15	US-10-093-463-139	Sequence 139, App
Sequence 139, App	20	2908	34.5	5175	15	US-10-085-198-37	Sequence 37, Appli
Sequence 37, Appli	21	2824.5	33.5	4727	12	US-10-332-447-56	Sequence 56, Appli
Sequence 56, Appli	22	2660	31.6	1548	10	US-09-822-846-97	Sequence 97, Appli
Sequence 97, Appli	23	1677	19.9	1346	9	US-09-562-436-337	Sequence 337, App
Sequence 337, App	24	1677	19.9	1346	9	US-09-880-107-3406	Sequence 3406, Ap
Sequence 3406, Ap	25	1677	19.9	1346	9	US-09-354-531-1033	Sequence 1033, Ap
Sequence 1033, Ap	26	1677	19.9	1346	14	US-10-171-581-333	Sequence 333, App
Sequence 333, App	27	1674	19.9	2645	15	US-10-094-749-1290	Sequence 1290, Ap
Sequence 1290, Ap	28	1553.5	18.4	6491	9	US-09-354-456-2118	Sequence 2118, Ap
Sequence 2118, Ap	29	1553.5	18.4	6491	12	US-10-342-887-559	Sequence 559, App
Sequence 559, App	30	1540	18.3	2512	15	US-10-108-260A-1795	Sequence 1795, Ap
Sequence 1795, Ap	31	1481	17.6	8040	15	US-10-191-803-169	Sequence 169, App
Sequence 169, App	32	1480	17.6	8350	14	US-10-072-900-3	Sequence 3, Appli
Sequence 3, Appli	33	1480	17.6	9112	14	US-10-072-900-1	Sequence 1, Appli
Sequence 1, Appli	34	1480	17.6	10365	14	US-10-098-939-1	Sequence 1, Appli
Sequence 1, Appli	35	1480	17.6	10381	14	US-10-098-939-2	Sequence 2, Appli
Sequence 2, Appli	36	1480	17.6	10423	14	US-10-098-939-3	Sequence 3, Appli
Sequence 3, Appli	37	1479	17.6	7860	12	US-10-617-334-2	Sequence 2, Appli
Sequence 2, Appli	38	1479	17.6	7860	15	US-10-452-510-2	Sequence 2, Appli
Sequence 2, Appli	39	1479	17.6	10412	14	US-10-313-641-7	Sequence 7, Appli
Sequence 7, Appli	40	1479	17.6	10412	15	US-10-428-551-7	Sequence 7, Appli
Sequence 7, Appli	41	1477.5	17.5	6786	14	US-10-313-641-8	Sequence 8, Appli
Sequence 8, Appli	42	1477.5	17.5	6786	15	US-10-428-551-8	Sequence 8, Appli
Sequence 8, Appli	43	1476	17.5	9870	10	US-09-984-827-103	Sequence 103, App
Sequence 103, App	44	1475	17.5	9497	16	US-10-429-160-3	Sequence 3, Appli
Sequence 3, Appli	45	1475	17.5	9741	9	US-09-846-456-10	Sequence 10, Appli

```
Db      3160 CAACATC-----:::---CCAGTGTTCAGGAGCTGAGCTTCGAGGT 3195
Qy      1323 llyslsGlyGluileuGlyLeuGlyProaenglyAlaGlyLysSerThrIleI 1343
Db      3196 GAAGAAGGGCAACCTTCGCTGTGGGAGCAGTGTGCTGGGAGAGTACAGTGT 3255
Qy      1343 eAniileLeuValGlyAspIleuProThrSerGlyGlnValPheLeuGlyAspTyrse 1363
Db      3256 ::::---CCAGTGTTCGAGCGCTTACAAACCCATGGCTGGACAGTGTCTTA----- 3304
Qy      1363 rSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAs 1383
Db      3305 -----GATGGCAAGAAATATA-----CAACTCAA 3330
Qy      1383 nProLeuTTPProAspThrThrLeuGlnGluHisPheGluileTyr----- 1398
Db      3331 CGTCCAGTGG-----CTCGCGCCACCTGGGCATTGTGCCAGGAGCCCAT 3378
Qy      1399 -----GlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSe 1414
Db      3379 CCGTGTGTGACTGCAGCATCACCGAAGACATCGCTACGGAGACAACAGCCGTGTCGTGC 3438
Qy      1414 r-----ArgileThrHisAlaLeuAspLeuLysGluHisLeuGlnLysTh 1429
Db      3439 TCATGAGGAGATCGTGGAGGCGCCAGGAGGCCACATCCACCAGTTTCATCGACTCACT 3498
Qy      1429 rValLysLys-----LeuProAlaGlyIlely 1438
Db      3499 GCCTGAGAATAACAACACCAGAGTGGGAGACAAGGGACTGCTCGGGCGGCAGAA 3558
Qy      1438 sArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspG1 1458
Db      3559 GCAGCGCATCGCCATCGCGCGCCCTCGTCAGACAGCCTCACATCTTCTGTGATGA 3618
Qy      1458 uProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAl 1478
Db      3619 AGCGACATCAGCTCTCGATACGGAGAGTGAAGAGTGTCCAGGAAGCGCTGGACAAAGC 3678
Qy      1478 aPheLysAsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluLalaGluAl 1498
Db      3679 C-----AGGGAAGGCGCACCTGCTATTGTATCGCGCACCGCCTGTCCACCATCGAA 3732
Qy      1498 aValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValG1 1518
Db      3733 C---GCAGACTTCATCGTGTGATTCAGAACGCCAGGTCAAGGAGCACGCCACCA 3789
Qy      1518 nHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAsp----- 1536
Db      3790 GCAGTGTGTCGCCCAAGAGGCATC---TATTTCGATGGTTCAGGCTGGAGCAAGCG 3846
Qy      1537 -----TrpIleGluAsnLeu-G 1542
Db      3847 CTCATGAGCTGGAGTATTAGTGCTCAAGTATTCTAATATTGGTGTCAACATGGC 3906
Qy      1542 luValAspArgLeuGlnArgGluileGlnTyrIlePheProAsnAlaSerArgGlnGluS 1562
Db      3907 ACGTAAACCAAGATTAAAGTTAAAGCAGCTGTAAAGTAAATTTCATCAAGACGAGAAG 3966
Qy      1562 erPheSerSertileLeu 1567
Db      3967 CCTTCAGAGACTTCATA 3983
```

Search completed: April 13, 2004, 02:07:49  
Job time : 426 secs

Qy	621	ArgLysLeuSerLeuGlyIleAlaValleuGlyAsnProLysIleLeuLeuLeuAspGlu	640
Db	1628	CAGAGGATCGCAATTCGCGGGCCCTGTCGCGAACCCCAAGATCTCTTTGGTGGATGAG	1687
Qy	641	ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLysTyrArg	660
Db	1688	GCCACGTCGAGCTTTGGACACAGAAGCGAAGCCGTGTTCAGGCCGCTCGATGAAGCCT	1747
Qy	661	LysAlaAsnArgValThrValPheSerThrHisPheMetAsp	677
Db	1748	AGAGAAGCCGACCAACCAATGTGATAGCTCACCGCTGTCTACAGTGCAGATGCTGAC	1807
Qy	678	IleLeuAla	695
Db	1808	GTCATTCCTGCTTTTGATGTGGTGCATGTGTGAGCAAGCA	1852
Qy	696	SerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAsp	715
Db	1953	CATGAAGAGCTCATGAAGAAG--GGCATTTACTTCAACTGTTCATG-----	1900
Qy	716	LysTyrCysAlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThr	735
Db	1901	-----ACACAGACTAGAGGAATGAATTGAA-----CCAGGAAT--	1936
Qy	736	LeuLeuGlnGlnAsnAspGlnLeuValTyrSerLeuProPheLysAspMetAspLys	755
Db	1936	-----	1936
Qy	756	PheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeuGlyVal	773
Db	1937	-----AATGCTTATGAATCCCAAGTGACACTGGTGCTCTGAGTTGACT	1981
Qy	774	TyrGlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGlu	793
Db	1982	TCGAATAAATCAAAATCTCTTTAATAGGAGATCAATTCGAGAGTATCCACAGAAGA	2041
Qy	794	IleAspGlnAlaAspTyrSerValPheThrGlnProLeuGluGluGluMetAspSer	813
Db	2042	CAAGACCAG--GAGAGAAGACTTAGTTCGAAGAGGATGTGGATGAAGATGTCCTATG	2098
Qy	814	LysSerPheAspGluMetGluGlnSerLeuLeuLeuSerGluThrLysAlaSerLeu	833
Db	2099	GTTTCCCTTTGGCAGATCCCTAAAG-----CTAAATATTAGTGAATCGCCCTATTAGTT	2152
Qy	834	ValSerThrSerLeuTrpLysGlnMetTyrThrIleAlaLysPheHis	851
Db	2153	GTGGGTGCTACTTTCGTGCTGT-TATAATGGGTGCTATACACAGGTGTTCCTATAGTGT	2211
Qy	852	-----PhePheThrLeuLysArg-GluSerLysSerValArgSerVal	865
Db	2212	TTCAAGAATGTAGGGGTTTTTTCAAGAGACGACGACCATGAACCAACCAACGGAATTG	2271
Qy	865	LLeuLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPheLeuValHisHis	885
Db	2272	TAACCTGTTTCCCTTCTCTTCTGTCATCGGGAATGATTCTTTTGTACGTAC	2326
Qy	885	rPheLysAsnAlaValValProLysLeuValProAspLeuTyrPheLeuLysProGlu	905
Db	2327	-----TTCTTTCAAGGCTTCACATTTGGCAAGCTGG	2358
Qy	905	YAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAspIle	925
Db	2359	AGATACCTCACCAAGGACTCCGATACATGGTCTTCAAATCCATGTCGCGACAGAGAT--	2416
Qy	925	eSerAspLeuLeuSerPhePhe	940
Db	2417	-----ATAAGCTGGTTTGATGACCATAAAAAACACCACTGGCTGCTGACTACCA	2466
Qy	940	rMetIleAsnAspSerAspTyrValSerValAlaProHisSer	956
Db	2467	GCTCGTAGTGACCCCTTCTAATGTTAAAGGGGCTATGGGCTCCAGGCTTCTGTAGTTAC	2526

Qy	956	aLeuAenValMetHisSerGluLysAspTyrValPheAlaValPheAsnSerThrMe	976
Dd	2527	CCAGAAATGTAGCAAACCTTGCGCACAGAATTATCTTACTCCTTAGCTCATGGCTGGCAGCT	2586
Qy	976	tValTySerLeuProIleuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAs	996
Dd	2587	TACACTTTTACTTGTTAGTAATATATACACTCAT-	2620
Qy	996	nValThrGluThrIleGlnIleTrpSerThrPropheGIngluileThrAspIleVa	1016
Dd	2620	-	2620
Qy	1016	lPheLysIleGluLeuTyrPheGlnAlaLaLeuLeuGlyIlelleValThrAlaMetPr	1036
Dd	2621	-	2638
Qy	1036	oProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysLeuAlaTyrThrGlnLeuLy	1056
Dd	2639	-	2688
Qy	1056	sLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValValAspIleProLe	1076
Dd	2689	GATCTCTGGGAATC-	2724
Qy	1076	uPhePheIleIleLeuLeuLeu-	1088
Dd	2725	CITTCGCACCTGTTGCTCTTGTACTAGGAGCAGAGAGTTTGAAACTATGTATGCCACAG	2784
Qy	1088	uLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaValPheCy	1108
Dd	2785	CITTCGATACCATAcCAGAAATGCTTTGAAGAAAGCACACACGCTTTGGGATCACCTTC-	2842
Qy	1108	sLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPheThrPhely	1128
Dd	2843	-	2895
Qy	1128	sLysIleLeuAsnThrLysGluPheTrpSerPhe-	1143
Dd	2896	TGCTACTTGGTGGCCAGACAACCTCATGACGTTTGAAAATGTTATGTTGTTATTTCTGC	2955
Qy	1143	lAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyrThrIleAl	1163
Dd	2956	TGTTGCTTTGGTGCCATGCACGACGAGGAATACCACTTCATTC-	2998
Qy	1163	aThrIleLeuHisTyr-AlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLe	1183
Dd	2999	-	3010
Qy	1183	uIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnPr	1203
Dd	3011	-	3055
Qy	1203	oTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePh	1223
Dd	3055	-	3055
Qy	1223	eLeuLeuGlnTyr-Tyr-GluLysLysTyrGlyGlyArgSerIleArgLysAspProPhePh	1243
Dd	3055	-	3055
Qy	1243	eArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAs	1263
Dd	3056	-	3064
Qy	1263	pGluApGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCy	1283
Dd	3065	-	3115
Qy	1283	sCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLy	1303
Dd	3116	-	3159
Qy	1303	sAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSerPheCysVa	1323

```

1429 rVAllySlyS-----LeuProAlaGlytley 1438
      ::::
3499 GCCTGAGAAATACAAACACACAGAGTGGGAGAGCAAAAGGAGCTCAGCTTCGCGCGGCGAGAA 3558
      ::::
1438 sAGSLySLeUCySPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGI 1458
      ::::
3559 GCAGCGCATCGCCATCGCGCGCGCCCTCGTCGACAGACGCTTACATCTTACTTCTGGATGA 3618
      ::::
1458 uProSerThrGlyMetAspProIlySAlaLySGlnHisMetTrpArgAlaIleArgThrAl 1478
      ::::
3619 AGCGACATCAGCTCTGTGATACGAGAGTGAAGAGGCTGCCAGGAAGCGCTGGACAAAGC 3678
      ::::
1478 aPheLySAsnArgLySArgAlaAlaIleLeuThrThrHisTyMetGluGluAlaGluAl 1498
      ::::
3679 C-----AGGGAAGCGCGCACCTGCATTGTGATCGCGCACCGCTGTCCACATCCAGAA 3732
      ::::
1498 aValCySAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGI 1518
      ::::
3733 C--GCAGACCTTGCATCGTGTGATTCAGAACGCCAGGTCAAGGAGCACGGCACCCACCA 3789
      ::::
1518 nHisLeuLySerLySPheGlyLySGlyTyTrpPheLeuGluIleLeuLySAsp----- 1536
      ::::
3790 GCAGCTGCTGCGCCAGAAAGGCATC---TATTTCGTGATGGTTTCAGCTCGAGCAAAAGCG 3846
      ::::
1537 -----TrrIleGluAsnLeu-G 1542
      ::::
3847 CTCATGAGCTGGGAGTATTTGAGGTGCTAAGTATTTCTAATATTTGGTGTCAACATGGC 3906
      ::::
1542 luValAspArgLeuGlnArgGluIleGlnTyrrIlePheProAsnAlaSerArgGlnGluS 1562
      ::::
3907 ACGTAAACCAAGTTAAAGGTTAAAGCACTGTTAAAGGTAAATTTCATCAAGCAGAGAAG 3966
      ::::
1562 erPheSerSerIleLeu 1567
      ::::
3967 CCTTCAGAGACTTCATA 3983
      ::::

RESULT 15
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; CURRENT FILING DATE: 1999-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:
Pred. No.: 5,16e-32 Length: 4233
Score: 407.50 Matches: 268
Percent Similarity: 35.54% Conservative: 232
Best Local Similarity: 19.05% Mismatches: 474
Query Match: 4.64% Indels: 434
DB: 3 Gaps: 52

US-10-090-458-5 (1-1642) x US-09-450-105-1 (1-4233)

QY 296 SerIleValIlePheLeuLeuPhePheLeuTyGlyLeuSerValPhePheAlaLeu 315
      ::::

```

Db	662	ACCTTGTAAATTTGGCCGTCAGCCCTCTTATTGGGTGTGCATCTGCCATGTGGGCAAG	721
Qy	316	MetLeuThrProLeuPheTysLys-----SerLysHisValGlyIleVal	330
Db	722	GTAAGCTTCAATTTACTAATAAAGAACTCCAGGCTTATGCGAAAGCTGGAGCAGATTGCC	781
Qy	331	Glu-----PhePheValThrValAlaPheGlyPheIleGlyLeuMetIle	345
Db	782	GAAGAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGA-----	823
Qy	346	IleLeuIleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCys	365
Db	823	-----	823
Qy	366	ThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGly-----	383
Db	824	-----GGCACAAGAAGAACTTGAAAGGTACAATAAAAAATTTAGAA	865
Qy	384	-----AlaSerPheSerAsnLeuThrAlaGlyProTyr	394
Db	866	GAAGCTAAAGAGTGTGCATAAAGAAAGCCATCAGGCCCAACATTTCCATAGGTATTGCC	925
Qy	395	ProLeuIleIle-----ThrIleIleMetLeu	403
Db	926	TACCTGTGTGTCATGCGTCTTATGCATGGCATTTCTGTGTATGGAGCTCTCTTGGTCTCTC	985
Qy	404	Thr-----LeuAsnSerIlePheTyrValIleLeuAlaVal	415
Db	986	TCAAAATGAATTTCTATTGGACAGTGCTTACCGTCTTCTTCTATTTATTATGGGGACT	1045
Qy	416	Tyr---LeuAspGlnValIleProGly-----GluPheGlyLeuArgArgSerSerLeu	432
Db	1046	TTCAAGTATTGCATTTAGCCCCAAACATAGAAGCCTTTGCAATGCAAGAGGGCGACGC	1105
Qy	433	TyrPheLeu-----LysProSerTyrTrpSerLysSerLysArgAsn	446
Db	1106	TATGAATCTTCAAGATAATTGATATGACCCAAAGCATTCACAGCTCTCTCAACCAAGGA	1165
Qy	447	TyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGluIleIleGlu	466
Db	1166	CACAAA-----CCAGACAGTATAATGGGAAATTTGGAATTT-----	1201
Qy	467	ProValSerSerGluPheValGlyLysGluAlaIleArgIleSerGlyIleGlnLysThr	486
Db	1202	-----AAAAATGTTTACTTCAAC	1219
Qy	487	TyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAspIleTyrGlu	506
Db	1220	TACCCATCAGCAAGTGAA---GTTAAGATCTTGAAGGGCCTCAACCTGGAAGGTGAAGAGC	1276
Qy	507	GlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIle	526
Db	1277	GGGCAGACGGTAGCCCTGGTGTGGACAGGTGGCTGTGGGAAAGACCACTGTCCAGCTG	1336
Qy	527	LeuCysGlyLeuCysProProSerAspGlyPheAlaSerIleTyrGlyHisArgValSer	546
Db	1337	CTGCAGAGGCTCTACGACCCCATAGAGGGCGAGGTGAGTATCGACGACAGGACATCAGG	1396
Qy	547	GluIleAspGluMetPheGluAlaArgLysMetIleGlyIleCysProGlnLeuAspIle	566
Db	1397	ACCATCAATGTAGGTAT---CTGCGGGAATCATTTGGGTGTGTAGTCAAGAACCCGCTG	1453
Qy	567	HisPheAspValLeuThrValGlu-----GluAsnLeuSerIleLeu	580
Db	1454	CTGTGTTGCCACCATGTTGCCGAAAACATTCGTATGGCCGAGAAAAACGTCAACCATGGAT	1513
Qy	581	AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu	600
Db	1514	GAGATAGAGAAAGCTGTCAAGGAAGCCCAATGCCCTATGACTTCAATCATGAAA-----	1567
Qy	601	AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlnLys	620
Db	1568	CCCCACAAAATTTAACAACCTGGTGTGGTGGAGAGAGGGCGCAGGTGATGCTGGGACAGAAA	1627

Db 1936 ----- 1936  
Qy 756 PheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeuGlyVal-----IleSer 773  
Db 1937 -----AATGCTTATGAATCCCAAGTGCACATGCTGCTGAGTTGACT 1981  
Qy 774 TyrGlyValSerMetThrThrLeuGluAspValPheLeuLeuLeuGluValGluAlaGlu 793  
Db 1982 TCAGAAAATCAAAATCTCTTATAAGGAGATCAATTCGAGAGATTCACAGAGA 2041  
Qy 794 IleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluGluMetAspSer 813  
Db 2042 CAAGACCAAG--GAGAGAGACTTAGTTCGAAAGAGGATGCTGAGATGCTGCTATG 2098  
Qy 814 LysSerPheAspGluMetGluGlnSerLeuLeuLeuLeuLeuSerGluThrLysAlaSerLeu 833  
Db 2099 GTTCTCTTTTGCAGATCCAAAG-----CTAAATATTAGTAATGCCCTATTAGTT 2152  
Qy 834 ValSerThrMetSerLeuTrpLysGlnGlnMetTyrThrIleAlaLysPheHis----- 851  
Db 2153 GTGGGTGACTTGTGCTGT-TATAATGGGTGCATACACAGGTTGTCATAGTGT 2211  
Qy 852 -----PhePheThrLeuLysArg-GluSerLysSerValArgSerVa 865  
Db 2212 TTCAAAAGATTGTAGGGTTTTTTCAGAGACGACGACCATGAAACCAACGGAATTG 2271  
Qy 865 IleLeuLeuLeuLeuLeuPhePheThrValGlnIlePheMetPheLeuValHisHis 885  
Db 2272 TAACTGTTCCTTCCTTCCTTCCTGTCATGGGAATGATTTCTTTGTACGTAC----- 2326  
Qy 885 rPheLysAsnAlaValProIleLysLeuValProAspLeuTyrPheLeuLysProGln 905  
Db 2327 -----TTCTTTCAAGGCTTCACATTTGGCMAAGCTGG 2358  
Qy 905 YAspLysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAspIle 925  
Db 2359 AGAGATCTCCACAGCGACTCCGATACATGCTTCAATCCATGCTGGCGAGAGAT-- 2416  
Qy 925 eSerAspLeuIleSerPhePhe-----ThrSerGlnAsnIleMetValTh 940  
Db 2417 -----ATAAGCTGTTTGTATGACCATAAAACACCACTGCTGCTGACTPACCAG 2466  
Qy 940 rMetIleAsnAspSerAspTyrValSerValAlaProHisSer-----AlaAl 956  
Db 2467 GCTCGCTAGTAGCGCTCTAAATGTTAAAGGGCTATGGCTCCAGGCTGCTGTAGTTAC 2526  
Qy 956 aLeuAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMe 976  
Db 2527 CCAGATGTAGCAAACTTGGCACAGGAATTATCTTATCTTAGTCTATGGCTGCACCT 2586  
Qy 976 tValTyrSerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAs 996  
Db 2587 TACACTTTTACTGTGTAGTAATATATACCACTCAT----- 2620  
Qy 996 nValThrGluThrIleGlnIleTrpSerThrProPheGlnGluIleThrAspIleVa 1016  
Db 2620 ----- 2620  
Qy 1016 lPheLysIleGluLeuTyrPheGlnAlaLeuLeuLeuGlyIleIleValThrAlaMetPr 1036  
Db 2621 -----GCTTGGGTGGAATTATT----- 2638  
Qy 1036 oProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLys 1056  
Db 2639 -----GAAATGAACCTGTTGCTGCTACGCTTGAAGGACAGAAAGACTAGA 2688  
Qy 1056 sLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLe 1076  
Db 2689 GATCTCTGGGAAGATC-----GCTCAGAGAAGCAATTGAAA 2724  
Qy 1076 uPhePheIleIleLeuLeuLeu-----MetLeuGlySerLe 1088  
Db 2725 CTTCCGCACTGTGTCTCTTGTACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2784

Qy 1088 uLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaValValPheCy 1108  
Db 2785 CTTGCAGATACCATACAGAAATGCTTTGAAGAAAGCACACGCTCTTTGGGATCACCTTC-- 2842  
Qy 1108 sLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPheThrPheLys 1128  
Db 2843 -----GCCTTCACCCAGCCCATGATTATTTTCTATGCTGCTGTTTCCGGTTCGG 2895  
Qy 1128 sLysIleLeuAsnThrLysGluPheTrpSerPhe-----IleTyrSerVa 1143  
Db 2896 TGCTACTTGGTGGCAGAGACTCATGACGTTTGAATAATGTTATGTTGTTATTTCTGTC 2955  
Qy 1143 lAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyrThrIleAl 1163  
Db 2956 TGTGTCTTTGGTGGCAGCAGCAGGGAATACCATGATCATTC----- 2998  
Qy 1163 aThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLe 1183  
Db 2999 -----GCTCCTGACTAC----- 3010  
Qy 1183 uIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnPr 1203  
Db 3011 -----GGAAGGCCAAAGTCTCAGCATCCCAATCATCAGGATCATTCAG----- 3055  
Qy 1203 oTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePh 1223  
Db 3055 ----- 3055  
Qy 1223 eLeuLeuGlnTyrTyrGluLysLysTyrGlyGlyArgSerIleArgLysAspProphePh 1243  
Db 3055 ----- 3055  
Qy 1243 eArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAs 1263  
Db 3056 -----AAATCCCC----- 3064  
Qy 1263 pGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCy 1283  
Db 3065 -GAGATTGACAGCTACAGCAGCGAGGCTTGAAGCCCTAATGTTGTTAGAAGGA----- 3115  
Qy 1283 sCysGluGluLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLy 1303  
Db 3116 -----AATGTGAATTTAATGGAGTCATGTTCAATATCCCAACCCGACC 3159  
Qy 1303 sAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyrIleSerPheCysVa 1323  
Db 3160 CAACATC-----CCAGTGCTTCAGGGACTGAGCTTCGAGT 3195  
Qy 1323 lLysLysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyAlaGlyLysSerThrIleL 1343  
Db 3196 GAAGAAGGGCCAAACCGCTTCGCTGTGGGACAGCAGTGCTGGGAAAGAGTACAGTGT 3255  
Qy 1343 eAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSe 1363  
Db 3256 CCAGCTGCTCGAGCGCTTCTACAAACCCATGGCTGGGAACAGTGTCTTA----- 3304  
Qy 1363 rSerGluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAs 1383  
Db 3305 -----GATGCCAAAGAAATAAAA-----CAACTCAA 3330  
Qy 1383 nProLeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyr----- 1398  
Db 3331 GTTCCAGTGG-----CTCGGCCCACTGGGCGATTTGTTCCAGAGAGGCCAT 3378  
Qy 1399 -----GlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSe 1414  
Db 3379 CTTGTTTGTAGTCAGCATCACCGAAGACATCGCTACGGAGACACACAGCGCTGCTGTC 3438  
Qy 1414 r-----ArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLysTh 1429  
Db 3439 TCATGAGGAGATCGTAGGGCCCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3498

/ NUMBER OF SEQUENCES: 2  
 / CORRESPONDENCE ADDRESS:  
 / ADDRESSES: SmithKline Beecham Corporation  
 / STREET: 709 Swedeland Road  
 / CITY: King of Prussia  
 / STATE: PA  
 / COUNTRY:  
 / ZIP: 19406  
 / COMPUTER READABLE FORM:  
 / MEDIUM TYPE: Diskette  
 / OPERATING SYSTEM: DOS  
 / SOFTWARE: FastSeq for Windows Version 2.0  
 / CURRENT APPLICATION DATA:  
 / APPLICATION NUMBER: US/09/120,513  
 / FILING DATE: 22-JUL-1998  
 / CLASSIFICATION:  
 / ATTORNEY/AGENT INFORMATION:  
 / NAME: King, William T  
 / REGISTRATION NUMBER: 30,954  
 / REFERENCE/DOCKET NUMBER: GP50008  
 / TELECOMMUNICATION INFORMATION:  
 / TELEPHONE: 610-270-5015  
 / TELEFAX: 610-270-5090  
 / TELEX:  
 / INFORMATION FOR SEQ ID NO: 1:  
 / SEQUENCE CHARACTERISTICS:  
 / LENGTH: 4233 base pairs  
 / TYPE: nucleic acid  
 / STRANDEDNESS: single  
 / TOPOLOGY: linear  
 / MOLECULE TYPE: Genomic DNA  
 / US-09-120-513-1

Alignment Scores:  
 Pred. No.: 5,16e-32 Length: 4233  
 Score: 407.50 Matches: 258  
 Percent Similarity: 35.54% Conservative: 232  
 Best Local Similarity: 19.05% Mismatches: 474  
 Query Match: 4.84% Indels: 434  
 DB: 3 Gaps: 52

US-10-090-458-5 (1-1642) x US-09-120-513-1 (1-4233)

Qy 296 SerIleValIlePheLeuPhePheLeuTyGlyLeuSerSerValPhePheAlaLeu 315  
 Db 662 ACCCTTGAATTTTGGCCCTCAGCCCTCTATTGGGTTGTCATCGCCATGTGGGCAAG 721  
 Qy 316 MetLeuThrProLeuPheLysLysLysLysLysLysLysLysLysLysLysLysLys 330  
 Db 722 GTACTGACTTCTACTAATAAGGAACCTCAGGCTTATGCGGAAGCTGAGCAGTTGCC 781  
 Qy 331 Glu-----PhePheValThrValAlaPheGlyPheLeuMetIle 345  
 Db 782 GAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGA----- 823  
 Qy 346 IleLeuIleGluSerPheProLysSerLeuValThrLeuPheSerProPheCysHisCys 365  
 Db 823 ----- 823  
 Qy 366 ThrPheValIleGlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGly----- 383  
 Db 824 -----GGACAAAGAAAGAACTTGAAGGTACAAATAAATAATTAGAA 865  
 Qy 384 -----AlaSerPheSerAsnLeuThrAlaGlyProTyr 394  
 Db 866 GAAGCTAAAGAGTTGGCATAAAGAACCATCAGCGCCAAACATTTCCTAGTATTGCC 925  
 Qy 395 ProLeuIleIle-----ThrIleIleMetLeu 403  
 Db 926 TACCTGTGTGCTATCGCTTATGCACTGGCATTCTGGTATGGGACCTCCCTGGTCCTC 985  
 Qy 404 Thr-----LeuAsnSerIlePheTyValLeuLeuAlaVal 415

Db 986 TCAATGAATATTCTATTGGACAAGTCTTACCGTCTTCTCTCTCTATTATTGGGACT 1045  
 Qy 416 Tyr---LeuAspGlnValIleProGly-----GluPheGlyLeuArgSerSerLeu 432  
 Db 1046 TTCAGTATTGGACATTTAGCCCAACATAGAACGCTTTGCCAAATGCAAGAGGGGAGCC 1105  
 Qy 433 TyrPheLeu-----LysProSerTyTrpSerLysSerLysArgAsn 446  
 Db 1106 TATGAATCTTCAAGATAATTGATATAGCCCAAGCATTTGACAGCTTCTCAACCAAGGA 1165  
 Qy 447 TyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSerPheSerGluIleLeuGlu 466  
 Db 1166 CACAAA-----CCAGACAGTATATGCGAAATTTGGAATTT----- 1201  
 Qy 467 ProValSerSerGluPheValGlyLysGluAlaIleArgIleSerGlyIleGlnLysThr 486  
 Db 1202 -----AAAAATGTTTACTTCAAC 1219  
 Qy 487 TyrArgLysLysGlyGluAsnValGluAlaLeuArgAsnLeuSerPheAspIleTyGlu 506  
 Db 1220 TACCATCAGCAAGTGAA---GTTAAGATCTTGAAGGCGCTCAACCTGAAGGTGAAGAGC 1276  
 Qy 507 GlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIle 526  
 Db 1277 GGGCAGACGGTAGCCTCTGGTTGGCAGAGTGGCTGTGGGAAAGACACACTGTCCAGCTG 1336  
 Qy 527 LeuCysGlyLeuCysProProSerAspGlyPheAlaSerIleTyGlyHisArgValSer 546  
 Db 1337 CTGCAGAGGCTCTACGACCCCATAGAGGCGGAGTCACTATCAGCAGCAGAGCATCAGG 1396  
 Qy 547 GluIleAspGluMetPheGluAlaArgLysMetIleGlyIleCysProGlnLeuAspIle 566  
 Db 1397 ACCATCAATGTGAGGTAT---CTGCGGGAATCATTTGGGTGTGAGTCAGGAACCCGIG 1453  
 Qy 567 HisPheAspValLeuThrValGlu-----GluAsnLeuSerIleLeu 580  
 Db 1454 CTGTTTCCACCCAGTATGCGGAAACATTCGTCATGCGCGGAGAAACGTCACCATGGAT 1513  
 Qy 581 AlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGluValGlnLysValLeuLeu 600  
 Db 1514 GAGATAGAGAAAGCTGTCAAGAGGACCAATCCCTATGACTTCATCATGAAA-----CTG 1567  
 Qy 601 AspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLys 620  
 Db 1568 CCCACAAATTTACACCTCTGTTGTGAGAGAGGCGCAGCTGATGGGGACAGAAA 1627  
 Qy 621 ArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAspGlu 640  
 Db 1628 CAGAGATCGCATTTGCCCGGCCCTGTGTCGCAACCCCAAGATCCTTTTGTGTGATGAG 1687  
 Qy 641 ProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLysTyArg 660  
 Db 1688 GCCACGTCACCTTGGACAGAAAGCGAGCGCTGTGTCAGCCGCTCTGGTAAAGGCT 1747  
 Qy 661 LysAlaAsnArgValThrValPheSerThrHisPheMetAsp-----GluAlaAsp 677  
 Db 1748 AGAAGGCGCGACCACTTGTGATAGTCACTCCCGTCTGTACAGTGGCGCAATGTCTGAC 1807  
 Qy 678 IleLeuAla-----AspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGly 695  
 Db 1808 GTCAATCTGTTGTTGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1852  
 Qy 696 SerSerMetPheLeuLysSerLysTyPheGlyTyArgLeuSerMetTyIleAsp 715  
 Db 1853 CATGAAGAGCTCATGAAGAAGAAG---GGCATTCTTCAAACTTCTCATG----- 1900  
 Qy 716 LysTyrCysAlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThr 735  
 Db 1901 -----ACACAGACTAGAGAAATGAATTGAA-----CCAGGAAT--- 1936  
 Qy 736 LeuLeuGlnGlnAsnAspGlnGlnLeuValTySerLeuProPheLysAspMetAspLys 755

Qy 942 eAsnAspSerAspTyrVal-----SerValAlaProHisSerAlaAlaLe 957  
 Db 11465 TAATGGTACCCATTTTAAAAATACTACCGAGTGGTCTAAGCGCATATAGCAGGAGCTC 11524  
 Qy 957 uAsnValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetVa 977  
 Db 11525 CATCATGCCAACAAAGTTTGGAGTATCGTTATTAGTTGGTTAGCGGTTATCTTAT 11584  
 Qy 977 lTyrSerLeuPro-----lLeuValAsnIlelleSerAsnTyrTyrLeu 993  
 Db 11585 TTGTTAGTTTGACAGTTTGGAGTCATCTGGTTTGAACAAACAGAGTGAATGATGAT 11644  
 Qy 993 rHisLeuAsnValThrGluThrIleGlnIleTrpSerThr-----ProphePheGlnI 1011  
 Db 11645 TGATTTTAAAGCAGAAATGTAATACCGTCATCTTTTACCATGGTTACACCT 11704  
 Qy 1011 uIleThrAspIleValPheLysIleGluLeuTyrPheGlnAlaAlaLeuLysIleI 1031  
 Db 11705 GATTTTACCTGTACT-----ACAGCTATTGTTGTTATTGT 11740  
 Qy 1031 eValThrAlaMetProProTyrPheAlaMetGluAsnAlaGluAsnHisLysIleLysAl 1051  
 Db 11741 TTATGGGCTAATGACCGCACTCTCTGGGCTGATATTACTGT-----GG 11788  
 Qy 1051 aTyrThrGlnLeu-----LysLe 1057  
 Db 11789 TTACTTAGAAGTATGGGTATAGTTTCCAAATTCATTCATTCGTTATTGTTGGAAATC 11848  
 Qy 1057 uSerGlyLeuLeuProSerAla-----TyrTrpIleGlyGlnAlaValValAspIlePr 1075  
 Db 11849 AGTTGAGCTAGAGTAGAGCTGGTCAATTCAGTTATGTAGCAATTAAGCAAGGAA 11908  
 Qy 1075 oLeuPhePheIleIleLeuLeu-----MetLeuGlySerLeuLe 1089  
 Db 11909 CTTGATATTTGATCAAGTATTGAATTTGCTCATTTAGAACCTTTTCAACTCTATT 11968  
 Qy 1089 uAlaPheHis-----TyrGlyLeu-----TyrPh 1097  
 Db 11969 AGCTATAGGAATTTAGGATTAATTTCAATTAAGTAAGTAATTAACATTTGATTTATGG 12028  
 Qy 1097 eTyrThrValLysPheLeuAlaValPheCysLeuIleGlyTyrValProSerValI 1117  
 Db 12029 ATATGCTGTAATTTTACTAACAGCTTCAATGCTCATCTTTATCTGATTCATTAGTGT 12088  
 Qy 1117 eLeuPheThrTyr-----lLeuAlaSerPheThrPheLy 1128  
 Db 12089 AGTATTTTGTGGTCAATAGCTATATTGGGTTGGGATTCGCTGAATCTTTACTATC 12148  
 Qy 1128 sLysIleLeuAsnThr-----LysGluPheTrpSerPheIleTyrSerValAlaAl 1145  
 Db 12149 TGCTTTGCTCTTGACAGGTTTAGGAGATGGTATCTGGCAATTTAT----- 12194  
 Qy 1145 aLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyrThrIleAlaThrI 1165  
 Db 12195 ----CCTTGTGCTGGGTTACTCGGCTA-----ATGGGT--ACCTTAATAAATCT 12238  
 Qy 1165 eLeuHisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLeuIleSe 1185  
 Db 12239 GTGGTATTACTCT-----GGCACAGCTATT 12265  
 Qy 1185 rPheIleLys-----lLeuSerTrpLysAsnValArgLysAsnValAspThrTyrAsnPr 1203  
 Db 12266 TTTTAAAGCAACAGCTTTTAAATTTGG----- 12290  
 Qy 1203 cTrpAspArgLeuSerValAlaVal-----lLeuSerProTyrLeuGlnCysValLe 1220  
 Db 12291 -----CTGGAGTCGCGAGTCCACTTAATGCGTTTAATCCTTAGTATAAT 12340  
 Qy 1220 uTrpIlePheLeuLeuGlnTyrTyrGluLysIleGlyArgSerIleAlaGlyAs 1240  
 Db 12341 TTGGTTC-----GACAGATGGCAAGACGATAG----- 12368

Qy 1240 pProPheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluProProAs 1260  
 Db 12368 ----- 12368  
 Qy 1260 pAsnGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeuMetGl 1280  
 Db 12369 -----AGTCATGAATAAAGGAAAGGAACTTTTCAACATG-ACCTATATTGG 12417  
 Qy 1280 yCysGlnCysCysGluLeuLysProSerIleMetValSerAsnLeuHisLysGluTyrAs 1300  
 Db 12418 T-----GTTAGTCATCTCAAAAAGTGTATAA 12444  
 Qy 1300 pAspLysLysAspPheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSe 1320  
 Db 12445 AACTCAGGAAGC-----CTCACTAAGGAGCGTTAAAGATATTAC 12486  
 Qy 1320 rPheCysValLysGlyGluIleLeuLysGluLeuGlyProAsnGlyAlaGlyLysSe 1340  
 Db 12487 GTTCTCAGTTCAAGAAGGGAATTTATGCTATTATGGGTGAATCTGGCTCAGGGAAGTC 12546  
 Qy 1340 rThrIleAsnIleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGl 1360  
 Db 12547 AACTCTCTTAATATCTAGCTTGTATGGATATCCAAAGTAGTGGTCATATCATCTTCAA 12606  
 Qy 1360 yAspTyrSerSerGluThrSerGluAspAspSerLeu-----LysCysMe 1376  
 Db 12607 TAACTATCAATTAGAGAAAGTTAAAGATGAAGAGGCTGCTGTTTTAGAGTCGGCATAT 12666  
 Qy 1376 tGlyTyrCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGl 1396  
 Db 12667 TGGTTTATTTTCAAAATTCATCTTTAAATATCTCAATATTAATAAGACAACTCTGTT 12726  
 Qy 1396 uIleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgI 1416  
 Db 12727 GATACAGATTATTATTTCGGGAAGTAAGGTGAATTCCTATGAAACAGATTACGTGATT 12786  
 Qy 1416 eThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGl 1436  
 Db 12787 AGCTGCTGTTGGTATGAATCTTGTCTATCTAAATATCCCTTATGAATATCTCGAGG 12846  
 Qy 1436 yIleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLe 1456  
 Db 12847 TCAACAAAGGTTAGCTATTGCCAGAGCTTTAATATGAATCCAGACTTGATATTGGC 12906  
 Qy 1456 uAspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleAr 1476  
 Db 12907 CGATGAGCCCAACAGGCAATTCGACTCTAAGACTTCTCAGCGAATCTTGAATTTGTTGC 12966  
 Qy 1476 gThrAlaPheLysAsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGluGluAl 1496  
 Db 12967 TAACATC-----AACGTAACGAAGACAATTCATAATGGTGACTCATAGCTCTAAAGC 13020  
 Qy 1496 aGluAlaValCysAspArgValAlaIleMetValSerGly-----GlnLeuAr 1512  
 Db 13021 TGCTAGTATGCAACCGAGTTCTTTTATCAAGGATGGTGTATTATTTCAATCAACTT-- 13078  
 Qy 1512 gCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyrPheLeuGlu 1531  
 Db 13079 -----GTTCTGGGTGTAAATCCAGGAGGAGC-----TTTTTAGAT 13114

RESULT 14  
 US-09-120-513-1  
 ; Sequence 1, Application US/09120513  
 ; Patent No. 6025160  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Brun, Kimberly  
 ; APPLICANT: Chenerly, Richard  
 ; APPLICANT: Ellens, Harma  
 ; APPLICANT: Field, John  
 ; APPLICANT: Yue, Lin  
 ; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE  
 ; SEQUENCES ENCODING RAT MDR1B2 AND  
 ; SCREENING METHODS THEREOF



9519 TTGCGAATAAGTAAGTATCTATCTCAAGAACCAAAATCTTTCTGGGACAATTGCT 9578  
340 -----PheIleGlyLeuMetIleI 346  
9579 GAACATTATTAGATTTTAATCAGCAACAGATACAGAACTAGTCTTTGGAAATGTTCTTGAT 9638  
346 e-----LeuIleGlySerPheProLysSerLeuValTrpLeuPheSe 360  
9639 ACTGTAGATATTACAGAACTTATTAGAAATTTACCGCGAAATTAGATTCTAAGGTG-- 9696  
360 rProPheCysHisCysThrPheValIleGlyLeuAlaGlnVal-----MetHis 376  
9697 -----ATGAATATGGTGAACCTCTCTGGAGGTCAAGATCCAGGATCTCA 9743  
376 sLeu---GluAspPheAsnGluGlyAlaSerPhe-----SerAsnLeuThrAlaG1 392  
9744 CTTCGACAGAGATTACTGAAGCAGGAGATGTTTATTATTAGATGAATCTT-- 9795  
392 yProTyProLeuIleLe-----ThrIleIleMe 402  
9796 -----TTGCCAATATTGATGAAGAGACTTGTCTTAAATAAAATTAATAATTTGCTGCTTAT 9851  
402 tLeuThrLeuAsnSerIlePheTyrValLeuLeuAlaValTyLeuAspGlnValIlePr 422  
9852 GCTGAATCACACAGCAAAATGCTTATGAAGTATTTCATAAT-CTAATAGATTAACCTCC 9910  
422 oGlyGlu-----PheGlyLeuArgSerSerLeuTyPheLeuLysProSerTyTr 440  
9911 CAGTAGTATCGTTTACCGATTGCTGATAAAAAATCTAGAAATTTTTCAGGAGCGGATTTTA 9970  
440 pSerLysSerLysAsnTyGlnGluLeuSerGluGlyAsnValAsnGly----- 457  
9971 ATAGAAAGTCGAGAAAT-----CTGAGTAAAGATCAGTTCTGTCGCGAAATY 10021  
457 ----- 457  
10022 AAATATTGTGATATATAAATAGCTTAAATCAATATCTCTAATAATTTGATTTTAAGCT 10081  
458 -----As 458  
10082 TTTTACTATTGTGAGTCTTTTACTCAAGATCTTTTGATTTTCTCGTAAAGTCTCTTAA 10141  
458 nIleSerPheSerGluIleIleGluProValSerSerGluPheValGlyLysGluAlaI 478  
10142 TTTGTTTTT-----ATACTAAAGCAGAAAGGAGGATATCATATGATGATATGCT 10195  
478 eaGlyLeSerGlyIleGlnLysThrTyArgLysLysGlyGluAsnValGluAlaLeuAr 498  
10196 AGAGACGAAATTTTAACATAACAGTTTGTGTAAGCAA-----ACAGCGGTTAA 10243  
498 gAsnLeuSerPheAspIleTyrglyGlnIleThrAlaLeuLeuGlyHisSerGlyTh 518  
10244 CCAATTGAATTTGAAAGTTGAACGTCATTCATTTATGTTGCTGGGCGCTAATGGTTC 10303  
518 rGlyLysSerThrLeuMetAsnIleLeuGlyLeuCysProProSerAspGlyPheAl 538  
10304 CGGCAATCAACACACTTAAATGATTAATCTGAATGCTTAAGAAAGACATCTGGTCACAT 10363  
538 aSerIleTyGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetI 558  
10364 TCATTATAGCGGACAGATTGAGCGCCAGGAT-----TTAGAAATAT 10408  
558 eGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSe 578  
10409 CGGGGCTCTGATTGAATCACCGCGCTTTTATGAAACCTGATCGCGCGTGAATAATTTAA 10468  
578 rIleLeuAsnSerIleLysGlyIleProAlaAsnIleIleGlnGluValGlnLysVa 598  
10469 GGTAAAGACCTTGATGCTGGTTTACCT---GATAGTCGATTTGATGAGGTTTAAAT 10525  
598 lLeuLeuAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyG1 618

10526 A-----GTGGATCTAACCAACACGGGTAAAAAAGAGCAGCGCAATTTTCTATGGG 10576  
618 yGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLe 638  
10577 CATGAAGCAGCGTCTGGTATTGCTATCGCACTTTTGAACCTCACCTCACTTTGATTC 10636  
638 uAspGluProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLy 658  
10637 GGATGAACCGCACTTAAGGACTTGATCTTATGGTATTACAGAGTTGCGTAATCTTATTCG 10696  
658 sTyArgLysAlaAsnArgValThrVal---PheSerThrHisPheMetAspGluAlaAs 677  
10697 TTTCTCTCCCTACACAAGGAATTACAGTTATTATTCAGTCATATCTTATCTGAGATTCA 10756  
677 pIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSe 697  
10757 GATGACAGCGGATCATATTGGTATCATGCTTAATGCGGACTG----- 10799  
697 rMetPheLeuLysSerLysTrpGlyIleGlyTyArgLeuSerMetTyIleAspLysTy 717  
10800 -----GGTTATCAG-----GATAGAAT 10816  
717 rCysAlaThrGluSerLeuSerSerLeuValLysGln----- 729  
10817 TCACCAAGATGAAGACTTGGAAAACTTTTACTGATGGTGTATGAGATACGAGAGG 10876  
730 -----HisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLe 744  
10877 TGAGTGATATGCTGGCGATGTTTCAGGCAGAAAGTT----- 10913  
744 uValTySerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeuAs 764  
10914 -----AAACTGAAGCGAAGTAGTTCGCAAGAAAGTTACTAGTTTTCG 10954  
764 pSerHisSerAsnLeuGlyValIleSerTyrglyValSerMetThrThrLeuGluAspVa 784  
10955 CCCCCATAAT-----ACATATTTATATGTTTATAGCACTGTGGGT 10999  
784 lPheLeuLys-----LeuGluValGluAlaGluIleAspGlnAlaAsp----- 798  
11000 ATTTAGTAAATAATGCTTATAATGTTGGTATGTCATGATTTTCCAGGCGCTCTAACCT 11059  
799 -----TySerValPheThrGlnGlnProLeuGluGluMetAspSerLysSe 815  
11060 TATTGTGCTGCTTAATAAATACTTACGAAGAAAAAGCTGCATTCATCGAGAGTTC 11119  
815 rPheAspGluMetGluGlnSerLeuLeuSerGluThrLysAlaSerLeuValSe 835  
11120 CTITGCCCATTTCTTTAAGAAAAATTTGGTTTGAATAATTTTATAACTGTTTATTATC 11179  
835 rThrMetSerLeuTrpLysGln---GlnMetTyThrIleAlaLysPheHisPheTh 854  
11180 TTAATTTTAGTAAATGAGTACTTTGGATAATACGATTAATCTGAAATCTTTTATTTC 11239  
854 r-----LeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeuLePh 872  
11240 CAATATTGGAAGA-----CTATCTTATCTGTTGGAG 11275  
872 ePheThrValGlnIlePheMetPheLeuValHisHisSerPheLysAsnAlaValPr 892  
11276 AATTAGCACTAGCTTCTTGGTTATATAGTATACACTTTCGCAAAATTCATTTGTC 11335  
892 oIleLysLeuValProAspLeuTyPheLeuLysProGlyAspLysProHisLysTyLy 912  
11336 TGTG-----GCTGACAAAGAAATCGGTTT-- 11360  
912 sThrSerLeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeuLeuSerPhe-- 931  
11361 -----TACCATACGTTGATAATTAATTAATGAGTAATTTTATTTCG 11404  
932 -----PheThrSerGlnAsnIleMetValThrMet----- 942  
11405 GAGTGTGTTTTCACACTCTTCCTGCTGGTGGCTTTGTCATATAGTTGGGAATCAGAT 11464

QY 1496 AlaGluAlaValCysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGly 1515  
 Db 1651 CCGGAA---TACTGGACGGGTGCGCATGCTCTCGCGTGGCGCTGATTGCACTGGAC 1707  
 QY 1516 ThrValGlnHisLeuLys 1521  
 Db 1708 ACGCCGACGGCTCAAG 1725

RESULT 13

US-09-627-376-3  
 ; Sequence 3, Application US/09627376  
 ; Patent No. 6342385  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OI, Fengxia Caulfield, Page Chen, Ping  
 ; TITLE OF INVENTION: MUTACIN I BIOSYNTHESIS GENES AND PROTEINS  
 ; FILE REFERENCE: UAB-17402/22  
 ; CURRENT APPLICATION NUMBER: US/09/627,376  
 ; CURRENT FILING DATE: 2001-05-30  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 3  
 ; LENGTH: 15567  
 ; TYPE: DNA  
 ; ORGANISM: Streptococcus mutans  
 US-09-627-376-3

Alignment Scores:  
 Pred. No.: 1,21e-31 Length: 15567  
 Score: 413.50 Matches: 363  
 Percent Similarity: 34.48% Conservative: 206  
 Best Local Similarity: 18.72% Mismatches: 617  
 Query Match: 4.91% Indels: 659  
 DB: 4 Gaps: 81

US-10-090-458-5 (1-1642) x US-09-627-376-3 (1-15567)

QY 17 LeuLeuLeuAsnTyrLeuIleLysCysArgThrLysSerSerValGlnGluIle 36  
 Db 7990 ATAATAATCAAGATCATCTTTA-----GTAATACTATA 8025  
 QY 37 LeuPheProLeuPhePheLeuPhePheLeuLeuIleSerMetMetHis 53  
 Db 8026 ATCAATATCAGGACAAATTTATTTGGCTCGTGGGTGGGAACAATTTACTAGACCAAT 8085  
 QY 54 -----ProAsnLysLysTyrGluGlu 60  
 Db 8086 ATGGTCAAGAGAGTGGTATATCAAGAACTTTTATGAAAGAACAAAGTCTTATCAACAA 8145  
 QY 61 Val-----ProAsnIleGluLeuAsnProMetAspLysPhe 72  
 Db 8146 GTAATAATGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8205  
 QY 73 -----ThrLeuSerAsnLeuLeuLeuGlyTyrThrProValThrAsnIleThrSerSerIle 91  
 Db 8206 GAGAGCTTCAGAAATTTATTA-----GTAATCACTATCTTCAAGGAGAT 8250  
 QY 92 MetGlnLysValSerThrAspHisLeuProAspValIleIleThrGluGluTyrThrAsn 111  
 Db 8251 CTTCAAAAGCTAGAGATATT-CTACTAATCAAGTT-----AATTATACTAGT 8297  
 QY 112 Glu-----LysGluMetLeuThr 117  
 Db 8298 GATGGTTATAAGAAATAGCACTTACAAAGCTTTGCGAGGTACTATAAATGAAGTGTTA 8357  
 QY 118 -----SerSerLeuSerLysProSerAsnPheValGlyValValPheLys 132  
 Db 8358 GAAGTTTCCAAATAGTAAAGAAAGAAATCTTTATCTTATTTGGTTGATATTTCA 8417  
 QY 133 AspSerMetSerTyrGluLeuArgPhePheProAspMetIleProValSerSerIleTyr 152  
 Db 8418 -----ATTATGACAGGCTTAATTTACTCTACGAATCACTAC 8453

QY 153 Met-----AspSerArgAlaGlyCysSerLysSerCysGluAlaAlaGln 167  
 Db 8454 TTACTTAAGAAATTTAGTTGACAGCAATCGTCTTTAATAATT--TGTTCTTTGTTCTTG 8511  
 QY 168 TyrTrpSerSerGlyPheThrValLeuGlnAlaSerIleAspAlaIleIleGlnLeu 187  
 Db 8512 TTTTGGGATTAGTTCTTTT-ATCATAGATGCTGGTTCACAGTATCTAATTTCAATTGATT 8570  
 QY 188 LysThrAsnVal-----SerLeuTrpLysGluLeu 197  
 Db 8571 GGTAATCAAGTAGTGTTTAACAGTCAAAATATATTTGGAAAAAATTTCTGATTGGACA 8630  
 QY 198 -----GluSerThrLysAlaVal-----IleMetGlyClnThrAlaValVal 211  
 Db 8631 GATAGTAAAGATGATTCTTCTGAAATGGCAGCCACCTTATTAAATGATAGTAGTGAATGATA 8690  
 QY 212 Glu-----IleAspThrPheProArgGlyVal----- 220  
 Db 8691 GAAATTTTATTAATTTCTACTTCTCAATCAATTAATTCAGTTATTGTTGATCAGGA 8750  
 QY 221 ----- 222  
 Db 8751 TCCTTAGTTATGCTATTGTTATTAATAGTAAAAATGCTTTTAGAGCTTATAGGATTGTC 8810  
 QY 223 IleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAla 237  
 Db 8811 TTGCTTTTATTTGTTTATGCAACCTTTTCTAGAAATTAAGCAAAATTAAGTAAAGAA 8870  
 QY 238 -----IleHisIleValAlaGluLysGluLysLysLys 249  
 Db 8871 ATCCAGGAGACAAAGCTGAACCTTATTAATATTTGCTCACAGTTGAGAGGACAAAGTCAA 8930  
 QY 250 GluPheLeuLysIleMetGlyLeuHisAspThrAlaPhe----- 262  
 Db 8931 ACA---ATAAAAGCTATAATGCTCAAGATTATGCTCTTTCAAAAATTTGATGAGCAAAAT 8987  
 QY 262 ----- 262  
 Db 8988 CGCCAATTTTCAAGATATCTTAATAGATAAAATTTTATAGCATTCTCTCTCTTT 9047  
 QY 263 -----TyrLeuSer----- 265  
 Db 9048 TTAATAATCTTAATCTTTTATGATTATAATTTGTTTGGCTAGGAATACAGAAAGTA 9107  
 QY 266 -----TyrValLeuLeuTyrThrSerLeu 273  
 Db 9108 CGTTACGAAATCTCACTAGTGTTCAGCAACTATTTTGTGTTTATGATGACACAAATTA 9167  
 QY 274 IlePheLeuMetSerLeuLeuMetAlaValIleAlaThrAlaSerLeuLeu----- 290  
 Db 9168 ATTAATCCAATTTGCAATTTATCAAAATTAGTTGCTCATATGGGATGCTTAATGGCGC 9227  
 QY 291 -----PheProGlnSerSerSerIle----- 297  
 Db 9228 GTGGAACGCTTTTGGAGTATATCAAGCTATTCAGAAAAAATGGAATCAAGAAAAAT 9287  
 QY 298 -----ValIlePhe-LeuLeuPhePheLeuTyrGlyLeuSerSerValPh 312  
 Db 9288 GATCAAAATTAATATCGGTTTGTATGATGTTTCTTATGCTTATGATAACCAAGAAAT 9347  
 QY 312 ePheAlaLeuMetLeuThrProLeuPheLysSerLysHisValGlyIleValGluPh 332  
 Db 9348 ATTATTGAAATGTGAATTTAACTTTTCAAAAAG-----GTACTTATTTCCATT 9398  
 QY 332 ePheValThrValAlaPheGly----- 339  
 Db 9399 GTTGGTGAAGAGTGGAGTTGGGAATCAACCTTACTTACTTTTAGACATAATATTATGTA 9458  
 QY 339 ----- 339  
 Db 9459 CCATCAAAAGGACCAATCTTAATAAAGGAATAGACTTAGAGAAATTAATTAAGACT 9518  
 QY 339 ----- 339

Db	747	-----	747	-----	1011
Qy	779	ThrThrLeuGluAspValPheLeuLysLeuGluValGluAlaGluIleAspGlnAlaAsp	798	PheIleTyrSerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMet	1158
Db	747	-----	747	-----	1011
Qy	799	TyrSerValPheThrGlnGlnProLeuGluGluMetAspSerLysSerPheAspGlu	818	GlyTyrThrIleAlaThrIleLeuHisTyrAlaPheCysIleIleIleProIleTyrPro	1178
Db	748	-----	748	-----	1011
Qy	819	MetGluGlnSerLeuLeuLeuSerGluThrLysAlaSerLeuValSerThrMetSer	838	LeuLeuGlyCysLeuIleSerPheIleLysIleSerTrpLysAsnValArgLysAsnVal	1198
Db	790	CTGGAAC-----	840	AspThrTyrAsnProTrpAspArgLeuSerValAlaValIleSerProTyrLeuGlnCys	1218
Qy	839	LeuTrpLysGlnGlnMetTyrThrIleAlaLysPheHisPheThrLeuLysArgGlu	858	ValLeuTrpIlePheLeuLeuGlnTyrTyrGluLysLysTyrGlyGlyArgSerIleArg	1238
Db	841	CTG-----	843	-----	1011
Qy	859	SerLysSerValArgSerValLeuLeuLeuLeuLeuPhePheThrValGlnIlePhe	878	LysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuProGluPro	1258
Db	843	-----	843	-----	1023
Qy	879	MetPheLeuValHisHisSerPheLysAsnAlaValValProIleLysLeuValProAsp	898	ProAspAsnGluAspGluAspGluAspValLysAlaGluArgLeuLysValLysGluLeu	1278
Db	843	-----	843	-----	1023
Qy	899	LeuTyrPheLeuLysProGlyAspLysProHisLysTyrLysThrSerLeuLeuGln	918	MetGlyCysGlnCysCysGluGluLysProSerIleMetValSerAsnLeuHisLysGlu	1298
Db	843	-----	843	-----	1056
Qy	919	AsnSerAlaAspSerAspIleSerAspLeuIleSerPhePheThrSerGlnAsnIleMet	938	TyrAspAspLysLysAspPheLeuLeuSerArgLysValLysLysValAlaThrLysTyr	1318
Db	844	-----	858	-----	1056
Qy	939	ValThrMetIleAsnAspSerAspTyrValSerValAlaProHisSerAlaAlaLeuAsn	958	Leu-----GlyAspTyrSerSerGluThrSerGluAspAspSerLeuLysCysMet	1376
Db	858	-----	858	-----	1296
Qy	959	ValMetHisSerGluLysAspTyrValPheAlaAlaValPheAsnSerThrMetValTyr	978	GlyTyrCysProGlnIleAsnProLeuTrpProAspThrThrLeuGlnGluHisPheGlu	1396
Db	858	-----	858	-----	1356
Qy	979	SerLeuProIleLeuValAsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThr	998	IleTyrGlyAlaValLysGlyMetSerAlaSerAspMetLysGluValIleSerArgIle	1416
Db	858	-----	858	-----	1416
Qy	999	GluThrIleGlnIleTrpSerThrProPhePheGlnGluIleThrAspIleValPheLys	1018	ThrHisAlaLeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGly	1436
Db	858	-----	858	-----	1476
Qy	1019	IleGluLeuTyrPheGlnAlaLeuLeuGlyIleIleValThrAlaMetProProTyr	1038	IleLysArgLysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeu	1456
Db	859	-----	888	-----	1536
Qy	1039	PheAlaMetGluAsnAlaGluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSer	1058	CACAAACAGCCCTAGCGCTGCCTGCCTATGCATCGGCCACCGTGCGTTCCTTG	1536
Db	889	GTACTCGCGAGGTCACCAACGAGCAAAATCCAGGCGCTGCGCGTCCGAGTGC	948	AspGluProSerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArg	1476
Qy	1059	GlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaValAlaAspIleProLeuPhePhe	1078	GATGACCTACTCTCGGCGTGGATCCGATCACCGCGCGAATCTTGGACCCACATCAAT	1596
Db	949	CAACTGGCTCCAGTCCCGCGGCTTCGAGAGCGCTTATATGAT-----	993	ThrAlaPheLysAsnArgLysArgAlaAlaIleLeu-----ThrThrHisTyrMetGluGlu	1495
Qy	1079	IleIleLeuLeuMetLeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyr	1098	GGCTTTG-----GACGCAAGGGGTGACCATCATGTCCACCGACATCTATGACGAA	1650
Db	993	-----	993	-----	1650
Qy	1099	ThrValLysPheLeuAlaValValPheCysLeuIleGlyTyrValProSerValIleLeu	1118		
Db	994	-----	1011	-----	
Qy	1119	PheThrTyrIleAlaSerPheThrPheLysLysIleLeuAsnThrLysGluPheTrpSer	1138		

1286 uLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLe 1306  
 1006 GAAACGGTCATTGAA-GCCAGTCGTTAAACGAAAAAGTTGGTGAT----- 1050  
 1306 uLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCysValLysLysG1 1326  
 1051 -----TTGCGCGCTACCGACCATGTCGATTTCCAGGTAAAGCGCGG 1091  
 1326 yGluIleLeuLysLeuGlyProAsnGlyAlaGlyLysSerThrIleAsnIleLe 1346  
 1092 TGAATTTTCGGCTCTTGGCCCAACGCGCGGAAATCCACCACCTTCAAAATGAT 1151  
 1346 uValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluTh 1366  
 1152 GTGCGCGCTCTGGTGTCCCACTCGGCAAGCGTGTGGGGAATGATCTCAAGT 1211  
 1366 rSerGluAspAspAspSerLysCysMetGlyTyrCysProGlnIleAsnProLeuTr 1386  
 1212 CAGC---TCTGGCAAGCGCGCCAGCATCTGGGCTATATGGCGCAAAATTTTCGCTGTA 1268  
 1386 pProAspThrLeuGlnGluHisPheGluLeuTyrGlyAlaValLysGlyMetSerAl 1406  
 1269 CGCAACCTCAGCGTCGACCAAACTCGCGCTTTTCTCCGCGCTATGGCGCTCGCGG 1328  
 1406 aSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLe 1426  
 1329 CGGCGCGCAAAATGAAAAATGCGCGGATGAGGATGCTTGGCTGAAAGTATCGC 1388  
 1426 uGlnLysThrValLysLeuProAlaGlyLysArgLysLysLeuCysPheAlaLeuSe 1446  
 1389 CGCCACGCGCGCGAGCTGCGCTGCTACAGCAGCGCTGCGCGCTGCGCTGCTC 1448  
 1446 rMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLy 1466  
 1449 CTGATGCGATGAACCGATATCTGTTTCTGATGAACCCACTCGCGCGCTGAGCCCT 1508  
 1466 sAlaLysGlnHisMetTyrAlaIleArgThrAlaPheLysAsnArgLysAlaAl 1486  
 1509 CACCGCGCGAGTCTCGCTACATATT---AATAGCATGTTGATAAAGCGTGACGCT 1565  
 1486 aLeuLeuThrHisTyrMetGluGluAlaValCysAspArgValAlaIleMe 1506  
 1566 AATGGTCACACCATTTATGATGAGCGGAA---TACTCGACCGCATCGGCTGCT 1622  
 1506 tValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLys 1523  
 1623 CTATCAGCGCAAGCTTATCGCAGCGGACCCCGGATGAGCTGAAAGCCAG 1674

RESULT 12  
 US-09-252-991A-1780  
 ; Sequence 1780, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; CURRENT INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 1780  
 ; LENGTH: 1824  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-1780

Alignment Scores: 3.79e-34 Length: 1824  
 Pred. No.:

Score: 421.50 Matches: 200  
 Percent Similarity: 28.64% Conservative: 111  
 Best Local Similarity: 18.42% Mismatches: 258  
 Query Match: 5.00% Indels: 517  
 DB: 4 Gaps: 22  
 US-10-090-458-5 (1-1642) x US-09-252-991A-1780 (1-1824)  
 QY 441 SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsnValAsnGlyAsnIleSer 460  
 Db 4 TCGAGGTCGATGCGAAAGCCGCGCACCATTCCTCCGGGGAGCGTTGAGATGCGAGTAAGC 63  
 QY 461 PheSerGluIleLeuGluProValSerSerGluPheValGlyLysGlu---AlaIleArg 479  
 Db 64 CTGCGCATCGCGCGCTGCTCGGGCGCG-----GGCGAAGACGCTGCGGTCGTC 114  
 QY 480 IleSerGlyIleGlnLysThrTyrArgLysGlyGluAsnValGluAlaLeuArgAsn 499  
 Db 115 ATCGAAGACGCTGACAAAGCATTTTC-----GGCGACGTAAAGGGCGCTGCGGGGC 162  
 QY 500 LeuSerPheAspIleTyrGluGlyGlnIleThrAlaLeuLeuGlyHisSerGlyThrGly 519  
 Db 163 TTGAGCGCGCATCTCACTATGCGCGCTACCGGCTGCTGCGCCCGCGCGCTGCG 222  
 QY 520 LysSerThrLeuMetAsnIleLeuCysGlyLeuCysProProSerAspGlyPheAlaSer 539  
 Db 223 AAGACGACGCTGATGCGGATTTCTGACCGGCTCTCTGGTCCGAAACGCGCGCGCTCACC 282  
 QY 540 IleTyrGlyHisArgValSerGluIleAspGluMetPheGluAlaArgLysMetIleGly 559  
 Db 283 TTGCAGCGCTATGCTGGTCAAGGACACGACGCCATTCATGTGCCACG-----GGC 336  
 QY 560 IleCysProGlnLeuAspIleHisPheAspValLeuThrValGluGluAsnLeuSerIle 579  
 Db 337 TACATCGCGCAACGCTTGGCTGTACGAAAGACCTGCTGCTGATGGAGAACATGCGCTG 396  
 QY 580 LeuAlaSerLysGlyIleProAlaAsnAsnIleIleGlnIleValGlnLysValLeu 599  
 Db 397 TATCGCAGTGTGGCTGCGACGACGCGCAACGCGAGCTGTTGCGCGAGCTGCTC 456  
 QY 600 LeuAspLeuAspMetGlnThrIleLysAspAsnGlnAlaLysLysLeuSerGlyGlyGln 619  
 Db 457 GACTTCAGCGGCTCGACCCCTTCCACAGCGCTCGCGCGACGCTCTCTGCGCGCATG 516  
 QY 620 LysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnProLysIleLeuLeuLeuAsp 639  
 Db 517 AAGCAAAATCTGGCGCTGCGCTGCGCTGCGCGCGCGCGCGCGCGCTGCTGCTGCTG 576  
 QY 640 GluProThrAlaGlyMetAspProCysSerArgHisIleValTrpAsnLeuLeuLys 658  
 Db 577 GAACCTGGAGTGGCGTGTGATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 636  
 QY 659 TyrArgLysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAlaAspIle 678  
 Db 637 CTGACCGATGAGGCGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 693  
 QY 679 LeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySerSerMet 698  
 Db 694 -----CGCTGC----- 699  
 QY 699 PheLeuLysSerLysTrpGlyIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCys 718  
 Db 699 ----- 699  
 QY 719 AlaThrGluSerLeuSerSerLeuValLysGlnHisIleProGlyAlaThrLeuLeuGln 738  
 Db 700 -----GAGACGCGCTGTTA 714  
 QY 739 GlnAsnAspGlnGlnLeuValTyrSerLeuProPheLysAspMetAspLysPheSerGly 758  
 Db 715 CTGAACAGGGGCGAGCTCTTGTGTTGATGCTCGG----- 747  
 QY 759 LeuPheSerAlaLeuAspSerHisSerAsnLeuGlyValIleSerTyrGlyValSerMet 778

```

Db      262 CACGCG-----GTGCTGGCTATATGCCGACAGAGTTTCGGCTCTGTATGAAGATCTGACG 315
Qy      573 ValGluGluAsnLeuSerIleLeuAlaSerIleLysGlyLeProAlaAsnAsnIleIle 592
Db      316 GTGATGGAACCTGACCTGACCGGACCTGCGCAGCGTCCACCGGC----- 363
Qy      593 GlnGluValGlnLysValLeuLeuAsp-----LeuAspMetGlnThrIle 607
Db      364 ---GAGGCGCGAAGAAATTTTCGACCGCTGCTGGAGTTCACCTCTCTTGGTCCATTC 420
Qy      608 LysAspAsnGlnAlaLysLysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIle 627
Db      421 ACCGAACGGCTGCGACGCAAGCTTTCCGCGGGATGAGCAAAAGCTGGCGCTCCCTGT 480
Qy      628 AlaValLeuGlyAsnProLysIleLeuLeuAspGluProThrAlaGlyMetAspPro 647
Db      481 ACCCTGGTGGCGATCCGAAGTACTGCTGCTCGATGACCGCGCTGGCGTCCGACCT 540
Qy      648 CysSerArgHisIleValTrpAsnLeuLeuLysTyrArgLysAlaAsnArgValThrVal 667
Db      541 ATCTCGCGCGGAGTATGCGAGATGGTGACGAGCTGGCGCGGCGGCGATGCTGATT 600
Qy      668 ---PheSerThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIle 686
Db      601 CTGTGAGACACCTCTATCTCGATGAGCGGACATGCGCGAC----- 645
Qy      687 SerGlnGlyMetLeuLysCysValGlySerSerMetPheLeuLysSerLysTrpGlyIle 706
Db      645 ----- 645
Qy      707 GlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerSerLeu 726
Db      645 ----- 645
Qy      727 ValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuValTyr 746
Db      646 -----GTGCTGCTGATGAACGAAGCAAACTGCTCTAC 678
Qy      747 SerLeuProPheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeuAspSerHis 766
Db      678 ----- 678
Qy      767 SerAsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeu 786
Db      678 ----- 678
Qy      787 LysLeuGluValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnPro 806
Db      679 -----CAGGGTGAGCGCGCGCTGACGCA----- 705
Qy      807 LeuGluGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuIleLeu 826
Db      706 -----ACCATGGCGCGGCACCTTT----- 726
Qy      827 SerGluThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGlnMetTyrThr 846
Db      727 -----CTGGTTTCCAG----- 738
Qy      847 IleAlaLysPheHisPheThrLeuLysArgGluSerLysSerValArgSerValLeu 866
Db      739 -----CAGCAGAGAACACCCGCGCTGCTGTG 765
Qy      867 LeuLeuLeuIlePheThrValGlnIlePheMetPheLeuValHisHisSerPhe 886
Db      765 ----- 765
Qy      887 LysAsnAlaValProIleLysLeuValProAspLeuTyrPheLeuLysProGlyAsp 906
Db      765 ----- 765
Qy      907 LysProHisLysTyrLysThrSerLeuLeuGlnAsnSerAlaAspSerAspIleSer 926

```

```

Db      766 -----CAGCGCGCGCTGAAGCTGCCGACAGTTTAGCAGCGGGGTATCCAG 810
Qy      927 AspLeuIleSerPheThrSerGlnAsnIleMetValThrMetIleAsnAspSerAsp 946
Db      811 -----GCCAATCGGTGGCTGTGATCTGATCTGAAAGAGCGCGCGC 849
Qy      947 TyrValSerValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyr 966
Db      850 ATTGAGGAGTCCAGCAGCATGCG----- 873
Qy      967 ValPheAlaAlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIle 986
Db      873 ----- 873
Qy      987 IleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTrpSerThr 1006
Db      873 ----- 873
Qy      1007 ProPheGlnGluIleThrAspIleValPheLysIleGluLeuTyrPheGlnAlaAla 1026
Db      873 ----- 873
Qy      1027 LeuLeuGlyIleIleValThrAlaMetProProTyrPheAlaMetGluAsnAlaGluAsn 1046
Db      874 -----GACATGCGCGCCCTGCGAGTGGCGCGCACCGCG----- 906
Qy      1047 HisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrp 1066
Db      906 ----- 906
Qy      1067 IleGlyGlnAlaValValAspIleProLeuPhePheIleIleLeuIleLeuMetLeuGly 1086
Db      907 -----CCGCGCTTTGAAGATGCGTTTATCGACCTGCTCGGC 942
Qy      1087 SerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaValVal 1106
Db      943 GGC----- 945
Qy      1107 PheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPheThr 1126
Db      945 ----- 945
Qy      1127 PheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyrSerValAlaAlaLeu 1146
Db      946 -----GCCGCGACCGCTGAA----- 960
Qy      1147 AlaCysIleAlaIleThrGluIleThrPhePheMetClyTyrThrIleAlaThrIleLeu 1166
Db      960 ----- 960
Qy      1167 HisTyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLeuIleSerPhe 1186
Db      960 ----- 960
Qy      1187 IleLysIleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnPro-TrpAspAr 1206
Db      961 -----TCGCGCTTGGCGGCG 975
Qy      1206 gLeuSerValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlu 1226
Db      976 ATTATCCATCGGTGACGCA----- 997
Qy      1226 nTyrTyrGluLysLysTyrGlyGlyArgSerIleArgLysAspProPhePheArgAsnLe 1246
Db      997 ----- 997
Qy      1246 uSerThrLysSerLysAsnArgLysLeuProGluProProAspAsnGluAspGluAspGlu 1266
Db      997 ----- 997
Qy      1266 uAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGlu 1286
Db      998 -----GCAAGAG 1005

```



QY 1616 nGluGluAspAsnSer 1622  
Db 1749 GGGGAAGGACGAGGACACC 1767

RESULT 10  
US-09-328-352-3466  
; Sequence 3466, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 3466  
; LENGTH: 1770  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-3466

Alignment Scores: 2,34e-38 Length: 1770  
Pred. No.: 461.00 Matches: 130  
Score: 30.18% Conservative: 129  
Percent Similarity: 17.98% Mismatches: 232  
Best Local Similarity: 5.47% Indels: 506  
Query Match: 4 Gaps: 25  
DB:

US-10-090-458-5 (1-1642) x US-09-328-352-3466 (1-1770)

QY 475 LysGluAlaIleArgIleGlnLysThrTyrArgLysLysGlyGluAsn--- 493  
Db 25 AAAGCAGTGGTTCAGCCCAATATTTATTCATGACTTTTAAAGCCGAAATATAAATTCA 84

QY 494 -----ValGluAlaLeuArgAsnLeuSerPheAspIleTyrGluGlnIleThrAla 511  
Db 85 GTAGATATTAGTCGCAATTTGTTGATTGAATGCAATTAATAAAGGAGAACTTACCGCT 144

QY 512 LeuLeuGlyHisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCys 531  
Db 145 TTATGCGTCCAGATGGTGGCGGTAAACAACTTTATTAAGGTTAATTCAGGATTATAT 204

QY 532 ProPheSerAspGlyPheAlaSerIleTyrGlyHisArgValSerGluLeuAspGluMet 551  
Db 205 AAAGCGAGTTCAGGTTCTCTTCATGTCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 264

QY 552 PheGluAlaArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeu 571  
Db 265 GTACAGACCGT-----ATTAGTTATATGCCCAACGTTTGGACTCTATGAGACTTA 318

QY 572 ThrValGluGlnAsnSerIleLeuAlaSerIleGlyIleGlyLeuProAlaAsnAsnIle 591  
Db 319 AGTGTAACAGAGATCTAAATTTATATGCGGCTTACACGGCGGTCCCAAAAATGACGT 378

QY 592 IleGlnGluValGlnLysValLeuLeuAspMetGlnThrIleLysAspAsnGln 611  
Db 379 AACCAAGATTAAACGTTTATGGAAATAACCGATCTTACTCAGTTTACTCAACGCTTA 438

QY 612 AlaLysLysLeuSerGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGly 631  
Db 439 GCAGGGCAACTTCTGCTGGTAATGAAGCAAAAGCTAGGATTGGCATGTACTCTAGTTGCT 498

QY 632 AsnProLysIleLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHis 651  
Db 499 TCCCAAGATTCTCTGCTGGTAATGAAGCAAAAGCTAGGATTGGCATGTACTCTAGTTGCT 558

QY 652 IleValTrpAsnLeuLeuLys-----TyrArgLysAlaAsnArgValThrValPheSer 669  
Db 559 GAUATTATGATCATTATGAACAACCTTGTTCAGGAAGAAATTTAAGTGTAATTATAAGT 618

QY 670 ThrHisPheMetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGly 689  
Db 619 ACGCTTATATGATGATGACAGAA----- 642

QY 690 MetLeuLysCysValGlySerMetPheLeuLysSerLysTrpGlyIleGlyTyrArg 709  
Db 643 -----AAGTGCCT----- 651

QY 710 LeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerSerLeuValLysGln 729  
Db 652 -----TACGTG-----TATATCATGCACGAA----- 672

QY 730 HisIleProGlyAlaThrIleLeuGlnGlnAsnAspGlnGlnLeuValTyrSerLeuPro 749  
Db 673 -----GGCAAAATTTAAGCAAGGTTACCTGACCAATTA----- 708

QY 750 PheLysAspMetAspLysPheSerGlyLeuPheSerAlaLeuAspSerHisSerAsnLeu 769  
Db 708 ----- 708

QY 770 GlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLysLeuGlu 789  
Db 708 ----- 708

QY 790 ValGluAlaGluIleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGlu 809  
Db 708 ----- 708

QY 810 GluMetAspSerLysSerPheAspGluMetGluGlnSerLeuLeuIleLeuSerGluThr 829  
Db 709 -----AAAGCGTTAGTACATGACGAACA 732

QY 830 LysAlaSerLeuValSerThrMetSerLeuTyrIleGlnGlnMetTyrThrIleAlaLys 849  
Db 733 -----TGGAA----- 738

QY 850 PheHisPheThrLeuLysArgGluSerLysSerValArgSerValLeuLeuLeuLeu 869  
Db 738 ----- 738

QY 870 LeuIlePheThrValGlnIlePheMetPheLeuValHisHisSerPheLysAsnAla 889  
Db 738 ----- 738

QY 890 ValValProIleLysLeuValProAspLeuTyrPheLeuLysProGlyAspLysProHis 909  
Db 739 -----ATMAAACCTTCAGAG----- 753

QY 910 LysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAspIleSerAspLeuIle 929  
Db 754 CAGATTAAAGACTCGGATTGTACAA-----GCCAGTTACTT 739

QY 930 SerPhePheThrSerGlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSer 949  
Db 790 GGA-----AATTCGGAGAGATTATTGAT 813

QY 950 ValAlaProHisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrValPheAla 969  
Db 814 GCAGTGCACAGAGAGAGCAAGTAACTTTATTAGTCGCCCAAAAGAACTA----- 864

QY 970 AlaValPheAsnSerThrMetValTyrSerLeuProIleLeuValAsnIleSerAsn 989  
Db 864 ----- 864

QY 990 TyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTrpSerThrProPhePhe 1009  
Db 865 -----TCAACAGATATTTTT 879

QY 1010 GlnGluIleThrAspIleValPheLysIleGluLeuTyrPheGlnAlaAlaLeuLeuGly 1029  
Db 880 CCT-----CTGGG 888

QY 1030 IleIleValThrAlaMetProProTyrPheAlaMetGluAsnAlaGluAsnHisLysIle 1049



```

; NAME/KEY: misc feature
; LOCATION: (1)...(2075)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-962

```

Alignment Scores:

Alignment Score:	7.19e-39	Length:	2075
Pred. No.:	467.00	Matches:	166
Score:	42.28%	Conservative:	116
Best Similarity:	24.98%	Mismatches:	236
Best Local Similarity:	5.54%	Indels:	150
Query Match:	4	Gaps:	20
DB:			

US-10-090-458-5 (1-1642) x US-09-833-381-962 (1-2075)

QY	1024	GlnAlaLeuLeuGly	-----llelleValThrAla	1034
Db	13	GAGGCTGCACTGATGGCTCTCGTGGTGGATGCTCTCGTCTCCATCTGTGTGGTCTTTGGC	72	
QY	1035	Met	-----ProProTyrPheAlaMetGluAsnAlaGluAsnHisIleLysAla	1051
Db	73	ATGTCTTTTTCGGCCGACCTTCACCTCTTGCTCCTCATTGAGGAGGAGTCACCGAGCC	132	
QY	1052	TyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyrTrpIleGlyGlnAlaVal	1071	
Db	133	AAGCACCTGCAGCTCATGGGGGGCTCTGCCACACCTCTACTGGGTGGCAACTTCTC	192	
QY	1072	ValAsp	-----IleProLeuPhePheIleIleLeuLeuLeuMetLeuGly	1086
Db	193	TGGGACATGTAACTACTTGGTGGCCAGCATGATCGTGGTGTCTATC	240	
QY	1087	SerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaVal	1106	
Db	241	---TTTCTGGCCTTCCACAGAGGGCATATGTGGCCCTGCCAACCTGCCTGCTCTCTG	297	
QY	1107	PheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPheThr	1126	
Db	298	CTGTGTCTACTACTGTATGGCTGGTGCATCACCGCTCATGTACCCAGGCTCTCTTCTTC	357	
QY	1127	PheLysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyrSerValAlaLeu	1146	
Db	358	TTTCTC	-----TCCGTGCCCAGCACCA	375
QY	1147	AlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyrThr	1161	
Db	376	GCCTATGTGTCTCACCTGCATAAACCTCTTATTGGCATCAATGAAGCATGGCCACC	435	
QY	1162	-----lleAlaThrIleLeuHisTyr	1168	
Db	436	TTTGTGCTTGAGCTTCTCTGTATCAGAAGTGCAGGAGGTGAGCCGGATCTTGAACAG	495	
QY	1169	AlaPheCysIleIleIleProlIleTyrProLeuLeuGlyCysIleIleSerPheIleLys	1188	
Db	496	GTCTTC---CTTATCTTCCCCCACTTCTGTGGCGCGGGGCTCATGTACGTGTCGG	552	
QY	1189	-----Ile	1189	
Db	553	AACCAGGCCATGGTGTATGCTTTTGAGGCTTGGGAGACAGGACGATCTCCAGTCAACCCCTG	612	
QY	1190	SerTrpLysAsnValArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSerVal	1209	
Db	613	CGTGGGAGTGGTGGCAAGACCTC	648	
QY	1210	AlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyrGlu	1229	
Db	649	GTGATACAGGGGCCCTCTTC	696	
QY	1230	LysLysTyrGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThrLys	1249	
Db	696	-----	696	
QY	1250	SerLysAsnArgLysLeuProGluPro	1264	

697 ---CGAAGCCAACCTCTCGCCACAGCCAGGGTGAGGTCTCTGCGCATCTCTGCGGAGGAGG 753

1265 AspGluuAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCys 1284

754 GACGAGGATGTACCCCGTGAACGGAGGGGGTG-----GTCCAAAGGAGCC 798

1285 GluGluLysProSerIleMetValSerAsnLeuHisLysGluLysTrpAspLysLysAsp 1304

799 ACCAGGGGATGTGTGGTCTGAGGAACCTTGACCAAGGTATAC----- 843

1305 PheLeuLeuSerArgLysValLysLysValAlaThrLysTrpLysSerPheCysValLys 1324

844 -----CGTGGCAGAGATGCCAGCTGTTGACCGCTGTGTCCCTGGGGATTCCTCC 891

1325 LysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValAlaGlyLysSerThrIleIleAsn 1344

892 CTGTGGTGGTGTTCGCTGTGGGGGCCAATGGAGCAGGGAAATAGTCGACGCTCCGC 951

1345 IleLeuValGlyAspIle-----GluProThrSerGlyGlnValPheLeu 1359

952 ATGCTAGCGGGCAGACATTTCTCCACGAGACCAACCTGTGCTCCACACCACTTATTG 1011

1360 GlyAspTyrSerSerGluThrSerGluuAspAspSerLeuLysCysMetGlyTyrCys 1379

1012 GGCACCTACTCCAGTGCACA-----CTCCTGTCTCTATTATTATTCTCTCT 1056

1380 ProGlnIleAsnProLeuTyrProAspThrThrLeuGlnGluHisPheGluIleTyrGly 1399

1057 TAAACCAATTACC---ATCTTTAACTCATGACGGCGCGGAGCACCTGGAGCTCTGGC 1113

1400 AlaValLysGlyMetSerAlaSerAspMetLysGluValLysSerArgIleThrHisAla 1419

1114 CGCTGCCCGGTGTCCGGAGGCGCAAGTTGCCACAGACCCCTGGCTCGGCGCTGGCGCGT 1173

1420 LeuAspLeuLysGluHisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArg 1439

1174 CTGGGACTCTCATGTATCGCAGACCGCGCTCGAGGCACCTACAGCGAGGAGAACAAAGC 1233

1440 LysLeuCysPheAlaLeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluPro 1459

1234 AAGCTGGCAGCGCCCTGGCGCTGTGGGTACCAAGCCGTGTGTCTCGACGAGCCG 1293

1460 SerThrGlyMetAspProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPhe 1479

1294 ACCACGCGATGACCCCAACGCGCGCGCTTCTTTGGAACAGCCCT---TTGCGCGTG 1350

1480 LysAsnArgLysArgAlaAlaIleLeuThrThrHisTyrMetGlu-GluAlaGluAlaVa 1499

1351 GTGCGGAAGGCGCTTCAGTGATGCTCACCTCCCATAGCATGGAGAGAGTGTGAAGGCT 1410

1499 LysAspArgValAlaIleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHis 1519

1411 CTGCTCGCGCTAGCATCATGTGTGATGGCGCGTTCCTGTGATGGCAGCAGCACAT 1470

1519 sLeuLysSerLysPheGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGln 1539

1471 CCTCAGCGAGATTCGCGCGGTTCACACTGACCCCTCGGTG----- 1515

1539 uAsnLeuGluValAspArgLeuGlnArgGluIleGlnTyrIle-----PheProAs 1556

1516 -----CCACCGCAAGGTCCCAGCGCGAGCGCGCTTCGTGCGCGCCGAGTTCCTCG 1568

1556 nAlaSerArgGlnGluSerPheSerSerIleLeuAlaTyrLysIleProLysGluuAspVa 1576

1569 GTCGGAGCTCGCGGAGGCAATGAGGCGCCCTGTGCGTTCACAGTGTCCGCGCGGAGGGCG 1628

1576 LginSerLysSerGlnSerPhePheLysLeuGluGluAlaLysHisAlaPheAlaIleGln 1596

1629 CTGCGCCCTGGCGCGCTCTTGGAGAGCTGGCGGTGCACCGCGCAGAGCAGCGCGTGGGA 1688

1596 uGluTyrSerPheSerGlnAlaThrLeuGluGlnValPheValGluLeuThrLysGluGln 1616

1689 GGACTTTTCGTGACGACAGCATGTCTGGAGGAGGTATCTCTTACTTCTTCCAGGACCA 1748

```

Db 2454 AAGCCAACTCTGCCACACCCAGGCTGAGGTCTCTGCCACTCTCTGGGAGGAGGACGA 2513
QY uAspValLysAlaGluArgLeuLysValLysGluLeuMetGlyCysGlnCysCysGluGlu 1286
Db 2514 GGATGTAGCCGTGACGCGGAGCGGGTG-----GTCCAAGGAGGACCCCA 2558
QY uLysProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLe 1306
Db 2559 GGGGGATGTGGTCTGAGGACTTGACCAAGGTATAC----- 2598
QY uLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCysValLysGlu 1326
Db 2599 -----CGTGCAGAGATGACCTCTTGACCGCTTGTGCTGGGATTCCTCCCTGG 2651
QY yGluLeuLeuGlyLeuGlyProAsnGlyAlaGlyLysSerThrIleLeuAsnIleLe 1346
Db 2652 TGAGTGTCTGGCTCTGGTGTGAATGAGGAGGAGAGCTCAGCTTCGCGATGGT 2711
QY uValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluTh 1366
Db 2712 GACGGGGACACATTGGCCAGCAGCGGCGAGGCTGTGCTGGCAGGCCAC---AGCGTGC 2768
QY rSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTr 1386
Db 2769 CCGGGAACCCAGTGTGCGCACCTCAGCATGGGATCTGCTCAATCCGATGCCATCTT 2828
QY pProAspThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAl 1406
Db 2829 TGAGTGTCTACGGCGCGAGCAGCCTGGAGCTGTCTGGCGGCTCGCGGGTCCGGA 2888
QY aSerAspMetLysGluValLysSerArgIleThrHisAlaLeuAspLeuLysGluHisLe 1426
Db 2889 GGCCAGGTTGCCACAGCCGCTGGCTGGCGCTGGCGCTGGGACTCTCATGGTACGC 2948
QY uGlnLeuThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAlaLeuSe 1446
Db 2949 AGACCGGCTTCAGGACCTCAGCGGAGGAGAACAAACGACGTGGCGACGCCCTGGC 3008
QY rMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProly 1466
Db 3009 GCTGTTGGGAGCCACCGCTGTGTTCTTGGACGAGCGACACACAGGATGACCCAG 3068
QY sAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaAl 1486
Db 3069 CGCGCGGCTTCTTTGGACAGCCTT---TTGGCGGTGGTGGGAGGCGGCTTCAGT 3125
QY aIleLeuThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMe 1506
Db 3126 GATGCTCACCTCCATAGCATGGAGGAGTGTGAAGCGCTCTGTCTGCGCGCTGGCCATCAT 3185
QY tValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLy 1526
Db 3186 GGTGAATGGCGGTTCCGCTGCTTGGCGACCCCGCAACATCTCAGGGCAGATTCGCGGC 3245
QY sGlyTyrPheLeuGluLeuLysLeuLysAspTrpIleGluAsnLeuGluValAspArgLe 1546
Db 3246 GGGTCACACACTGACCTCGGG-----GTGCCCGCCCAAGGTC 3284
QY uGlnArgGluIleGlnTyrIle-----PheProAsnAlaSerArgGlnGluSerPh 1563
Db 3285 CCACCGCGAGCGGCTTCGTGGCGGCGGAGTCCCTGGGCGGAGCTGCGCGAGGACCA 3344
QY eSerSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPh 1583
Db 3345 TGGAGGCGGCTGCGCTTCCAGCTGCGCGGAGGCGGCTGCGCGCTGGCGCGCTT 3404
QY ePheLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAl 1603
Db 3405 TGGAGAGCTGCGGTGACCGCGCAGACGCGGCTGGAGGACTTTTCC----- 3453
QY aThrLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAspAsnSer 1622
Db 3454 -----GTATCTTGTACTTCTCCAGGACACCGGAGGAGGAGGACGAGCACCC 3498

```

RESULT 8

US-09-621-976-409  
; Sequence 409, Application US/09621976  
; Patent No. 6639063  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Jobert, S.  
; APPLICANT: Giordano, J.Y.  
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
; FILE REFERENCE: GENSET 054PR2  
; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 409  
; LENGTH: 485  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 185..484  
US-09-621-976-409

Alignment Scores: 4.44e-45 Length: 485  
Pred. No.: 515.00 Matches: 100  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 6.11% Gaps: 4  
DB:

US-10-090-458-5 (1-1642) x US-09-621-976-409 (1-485)

```

QY 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20
Db 185 ATGTCCACTGCAATTAGGAGGTAGGAGTTGGAGACAGACCAACACTTCTACTGAAG 244
QY 21 AsnTyrLeuIleLysCysArgThrLysLysSerValGlnGluLeuLeuPheProLeu 40
Db 245 AATTCTTAATAATGACAGNACCAAAAGAGTAGTCTTCAGGAAATCTTTTCCACTA 304
QY 41 PhePheLeuPheTrpLeuLeuLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60
Db 305 TTTTCTTTTATTTTGTGTTAATATTAATAGCATGATGATCAATAAGAAATATGAAGAA 364
QY 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80
Db 365 GTGCCATAATAGAACCAATCTCTATGACACAGTTTACTCTTCTTAATCTAATCTTGA 424
QY 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100
Db 425 TATATCTCCAGTGACTAATATTAATTAAGCAGCATCATGAGAAAGTGTCTACTGATCATCTA 484

```

RESULT 9

US-09-833-381-962  
; Sequence 962, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 962  
; LENGTH: 2075  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:



Db 816 CTCTTTGTTAACTATAAGCTGCCGTGGCAGCGTTTACCTCTTATCAGACGACCTTTTCA 875  
Qy 1584 eLysLeuGluGluAlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaLh 1604  
Db 876 CAAATTAGAGCAGTCAGCAATAACCTTTAACTGGGAAGAAATACAGCTTTTCAGTGCAC 935  
Qy 1604 rLeuGluGlnValPheValGluLeuThrLysGluGlnGluGluAspAsnSerCysG 1624  
Db 936 ACTGGAGAGGTATTCTTAGAGCTTTCTAAAGAACAG---GAAGTAGGAAATTTTGATGA 992  
Qy 1624 yThrLeuAsnSerThrLeuTrpTrrGlu 1633  
Db 993 AGAAATTGATACAAACATGAGATGGAAA 1020

RESULT 7

US-09-833-381-958  
; Sequence 958, Application US/09833381  
; Patent NO. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 958  
; LENGTH: 3609  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-958

Alignment Scores:  
Pred. No.: 1,62e-57 Length: 3609  
Score: 647.00 Matches: 269  
Percent Similarity: 37.34% Conservative: 194  
Best Local Similarity: 21.69% Mismatches: 420  
Query Match: 7.68% Indels: 357  
DB: 4 Gaps: 38

US-10-090-458-5 (1-1642) x US-09-833-381-958 (1-3609)

Qy 639 AspGluProThrAlaGlyMetAppProCysSerArg-----HisIleVal-TrpAsnLe 656  
Db 84 GAACAAGCTACAGCGCGGCTCCGTCCTCCCAAGATCAACCGCTCCATGCGAAGCTGGACCA 143  
Qy 656 uLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPheMetAspGluAl 676  
Db 144 GCTAGAA---AACCTGTCCAACTTCATCAAGGCCATGGTCAGCTACGGCATGAACCTGT 200  
Qy 676 aspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLysCysValGlySe 696  
Db 201 GGACCTGTTCGAGGCCAACAGCCTGTTTGAGAGTGGGAACATGACGCGAGTGCAGGTGTC 260  
Qy 696 rSerMetPheLeuLysSerLys-----TrpG 705  
Db 261 TCTTCTCCCTTCGGCGGAGAGGTACCAATGCTCTTCCCGAGGAGACTGGGGTGGG 320  
Qy 705 yIleGlyTyrArgLeuSerMetTyrIleAspLysTyrCysAlaThrGluSerLeuSerSe 725  
Db 321 CQTGAGCGGGGCTCCCTGAAGCACACCCCTTTTGTCCACAGGCGACTCTCAGCTGCTGGC 380  
Qy 725 rLeuValLysGlnHisIleProGlyAlaThrLeuLeuGlnGlnAsnAspGlnGlnLeuVa 745  
Db 381 CTGTGTACAGACTGGGTGCCGGGACGCGCTGGTGGAGGAGCTGCCACACGAGCTGGT 440  
Qy 745 lTyrSerLeuProPheLysAspMet-----AspLysPheSerGlyLeuPheSerAlaLe 763  
Db 441 GCTGGTGTGCTCCCTACACGGGTGGCCCATGATGACGGCAGCTTCGCCACACTCTTCCGAGCT 500

Db 1285 GluGluValProSerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAsp 1304  
Db 6 GATGAGAACTGTTATAATTCGACGCTGTCTACACAAAGATATGCGAGGCCAGAGAAA 65  
Qy 1305 PheLeuLeuSerArgLysValLysValAlaThrLysTyrIleSerPheCysValLys 1324  
Db 66 AGTTGCTTTTCAAGGAGAGAGAAATAGCAGCAAGAAATATCTCTTCTGTGTTCAA 125  
Qy 1325 LysGlyGluIleLeuGlyLeuLeuGlyProAsnGlyValAspLysSerThrIleLeuAsn 1344  
Db 126 GAGATGAAATTTGGATTGCTAGGACCCCAATGGTGTGTAAGAAAGTTCACTATTAGA 185  
Qy 1345 IleLeuValGlyAspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSer 1364  
Db 186 ATGATATCTGGATCAACAAGCAACTGCTGGAGAGGTGGAACCTGAAGCTGCAGTTCA 245  
Qy 1365 GluThrSerGluAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnPro 1384  
Db 246 -----GTTTGGGCCACTGGGTGCTGCTCCCTCAAGAGAACGTCG 284  
Qy 1385 LeuTrpProAspThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMet 1404  
Db 285 CTGTGGCCCATGCTGACGTTGAGGAAACACCTGGAGGTGTATGCTGCCGTCAAGGGCTTC 344  
Qy 1405 SerAlaSerAspMetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGlu 1424  
Db 345 AGGAAGCGGACGGAGGCTCGCATCGCAAGATTAGTGTGCTTCAACTGCGATGAG 404  
Qy 1425 HisLeuGlnLysThrValLysLysLeuProAlaGlyIleLysArgLysLeuCysPheAla 1444  
Db 405 CAGCTGAATGTTCTGTGCAGAAATTAACACGAGGAATCACGAGAAAGTTGTGTTTGTG 464  
Qy 1445 LeuSerMetLeuGlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAsp 1464  
Db 465 CTGAGCCTCTGGGAAACTCCCTTCGTCGCTCTCGATGAACCATCTACGGGCTAGAC 524  
Qy 1465 -ProLysAlaLysGlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLysAr 1484  
Db 525 CCCACAGGCGACCCAGCAAAATGTGGCAGGCAATCCAGGCAAGTGTGTTAAAAACACAGAG 584  
Qy 1484 gAlaAlaIleLeuThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAl 1504  
Db 585 AGGTGCTCTGTGACCCCAATACCTGGTGGTGGCGGAGAGCTTGTGTGACCGTGTGGC 644  
Qy 1504 alleMetValSerGlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPh 1524  
Db 645 CATCATGTTGCTGGAAGGCTTAGATGCAATGGCTCCATCAACACCTGAAAAACAAACT 704  
Qy 1524 eGlyLysGlyTyrPheLeuGluIleLysLeuLysAspTrpIleGluAsnLeuGluValAs 1544  
Db 705 TGGCAAGGATTACATTCTAGAGCTTAAAGTGAAG-----GAAACGTCTCAAGTGAC 755  
Qy 1544 parGLeuGlnArgGluIleGlnTyrIlePheProAsnAlaSerArgGlnGlnSerPheSe 1564  
Db 756 TTTGTTCCACACTGAGATTCTGAGCTTTTCCACAGGCTGCAGGGCAGGAAAGGTATTC 815  
Qy 1564 rSerIleLeuAlaTyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePh 1584

US-10-090-458-5 (1-1642) x US-09-833-381-1082 (1-1350)

Alignment Scores:  
Pred. No.: 3,87e-85 Length: 1350  
Score: 900.50 Matches: 186  
Percent Similarity: 69.14% Conservative: 56  
Best Local Similarity: 53.14% Mismatches: 96  
Query Match: 10.69% Indels: 12  
DB: 4 Gaps: 3

Db 6135 GTTGGCAAGTTGGTGGAGTGGCGATTCGGAACCTGGGCTCGTGAAGTATGAGCAAAA 6194  
Qy 1429 ThrValLysLeuProAlaGlyLeuLysLeuCysPheAlaLeuSerMetLeu 1448  
Db 6195 TATGCTGGTAACATATAGTGGAGCAACAAACGAAAGCTCTTACAGCCATGGCTTTGATC 6254  
Qy 1449 GlyAsnProGlnIleThrLeuLeuAspGluProSerThrGlyMetAspProLysAlaLys 1468  
Db 6255 GCGGGCTCTCGTGGTGGTCTTGGATGAACCCACACAGGATGGATCCCAAGCCCG 6314  
Qy 1469 GlnHisMetTrpArgAlaIleArgThrAlaPheLysAsnArgLysArgAlaIleLeu 1488  
Db 6315 CGGTTCTTGGAAATGTGCCCTAAGTGTGTCAAG---GAGGGGAGATCATGTAGTGTCT 6371  
Qy 1489 ThrThrHisTyrMetGluGluAlaGluAlaValCysAspArgValAlaIleMetValSer 1508  
Db 6372 ACATCTCATAGTATGGAAGATGTGAGCTCTTGCATAGATGGCATCATGTGCAAT 6431  
Qy 1509 GlyGlnLeuArgCysIleGlyThrValGlnHisLeuLysSerLysPheGlyLysGlyTyr 1528  
Db 6432 GGAAGCTTCAGGTGCTTGGCAGTGTCCAGCATCTAAAAAATAGGTTTGGAGATGGTTAT 6491  
Qy 1529 PheLeuGluLysLeuLysAspTrpIleGluAsnLeuGluValAspArgLeuGlnArg 1548  
Db 6492 ACAATAGTTGTAGCAATAGCAGG-----TCCACCCCGAGCTGAAGCTGTCCAGGAT 6545  
Qy 1549 GluIleGlnTyrIlePheProAsnAlaSerArgGlnGlnSerPheSerSerIleLeuAla 1568  
Db 6546 TTCTTTGGACTTCATTTCTTGGAGTGTCTTAAAGAGAAACACCGGAACATGTCTACAA 6605  
Qy 1569 TyrLysIleProLysGluAspValGlnSerLeuSerGlnSerPhePheLysLeuGlu 1588  
Db 6606 TACCAGCTTCCA---TCCTCATATCTTCTGCGCCAGGATATTCAGCATCTCTCCAG 6662  
Qy 1589 AlaLysHisAlaPheAlaIleGluGluTyrSerPheSerGlnAlaThrLeuGluGlnVal 1608  
Db 6663 AGCAAAACGACTCCACATAGAGACTACTCTGTTTCTCAGACACACTTGCACCAAGTA 6722  
Qy 1609 PheValGluLeuThrLysGluGlnGluGluAspAsn----- 1621  
Db 6723 TTGTGAACCTTCCCAAGCAACCAAGTGTATGATGACCACTTAAAGACCTCTCATTTACAC 6782  
Qy 1622 -----SerCysGlyThrLeuAsnSerThrLeuTrpGluArgThr 1635  
Db 6783 AAAACACAGACAGTATGAGCTGTGAGTGTCTCAGTCTTCACATCTTTCTACAGGATGAGAAAGTG 6842  
Qy 1636 GlnGluAsp-----ArgValValPhe 1642  
Db 6843 AAAGAAAGCTATGTATGAGAAATCTCTGTT 6872

RESULT 5  
US-09-833-381-1085  
; Sequence 1085, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1085  
; LENGTH: 974  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-833-381-1085  
Alignment Scores: 3.24e-122 Length: 974  
Pred. No.: 1248.00 Matches: 245  
Score:

Percent Similarity: 100.00% Conservative: 2  
Best Local Similarity: 99.19% Mismatches: 0  
Query Match: 14.81% Indels: 0  
DB: Gaps: 0  
US-10-090-458-5 (1-1642) x US-09-833-381-1085 (1-974)  
Qy 1 MetSerThrAlaIleArgGluValGlyValTrpArgGlnThrArgThrLeuLeuLys 20  
Db 233 ATGTCCACTGCAATTAGGAGGTAGGAGTTTGAGACAGACACCAAGAACACTTCTTACTGAAG 292  
Qy 21 AsnTyrLeuLysCysArgThrLysSerSerValGlnGluLeuLeuPheProLeu 40  
Db 293 AATTACTTAATTAATGACACCAAAAGAGTAGTGTTCCAGAAATCTTTTCCACTA 352  
Qy 41 PhePheLeuPheTrpLeuIleLeuIleSerMetMetHisProAsnLysLysTyrGluGlu 60  
Db 353 TTTTITTTTITTTGGTTAATTAATTAGCATGATGATCCAAATAAGAAATAGAGAA 412  
Qy 61 ValProAsnIleGluLeuAsnProMetAspLysPheThrLeuSerAsnLeuLeuGly 80  
Db 413 GTGCTTAATATAGAACTCAATCTCTATGACAAAGTACTCTTCTTAATCTAATCTTTGGA 472  
Qy 81 TyrThrProValThrAsnIleThrSerSerIleMetGlnLysValSerThrAspHisLeu 100  
Db 473 TATTAATCCAGTCAATATATACAGCAGCATCATGCAAGAGTCTTACTGATCATCTA 532  
Qy 101 ProAspValIleIleThrGluGluTyrThrAsnGluLysGluMetLeuThrSerSerLeu 120  
Db 533 CCTGATGCTAATTAATGAGAAATATACAAATGAAAGAAATGTTTAAACATCCAGTCTC 592  
Qy 121 SerLysProSerAsnPheValGlyValValPheLysAspSerMetSerTyrGluLeuArg 140  
Db 593 TCTAAGCCGAGCACTTTGAGTGTGTTTCAAGACTCCATGCTCTATGACTTCTGT 652  
Qy 141 PhePheProAspMetIleProValSerSerIleTyrMetAspSerArgAlaGlyCysSer 160  
Db 653 TTTTCTCTGATATGATTCAGTATCTTCTATTATATGATGATCAAGAGCTGGTGTCTCA 712  
Qy 161 LysSerCysGluAlaAlaGlnTyrTrpSerSerClyPheThrValLeuGlnAlaSerIle 180  
Db 713 AAATCATGTGAGGTGTCTCAGTACTGTCTCTCAGGTTTTCAGTCTTTTACAGCATCCATA 772  
Qy 181 AspAlaAlaIleIleGlnLeuLysThrAsnValSerLeuTrpLysGluLeuGluSerThr 200  
Db 773 GATCTGCCATTATACAGTTGAGACCAATGTTTCTTTTGGAGGAGTGGAGTCAACT 832  
Qy 201 LysAlaValIleMetGlyGluThrAlaValValGluLeuLeuAspThrPheProArgGlyVal 220  
Db 833 AAAGCTGTTATTATGGGAGAAACTGCTGTTGTAGAAATAGATACCTTTCCCGGAGGAGTA 892  
Qy 221 IleLeuIleTyrLeuValIleAlaPheSerProPheGlyTyrPheLeuAlaIleHisIle 240  
Db 893 ATTTTATATACCTAGTTATAGCATTTTCCACTTTTGATACATTTTGGCAATTCATATC 952  
Qy 241 ValAlaGluLysGluLysLys 247  
Db 953 GTAGCAAAAAAAGAAAAA 973  
RESULT 6  
US-09-833-381-1082  
; Sequence 1082, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0

QY 882 ----- 882  
 Db 4176 GTGCCACCTTTGGCAAGTACCCAGCCTGGAACTTCAGCCTGGATGTACAAAGAACAG 4235  
 QY 882 ----- 882  
 Db 4236 TACACATTGTGACCAATGATGCTCTCTGAGACACGGGAACCTGGAACTCTTAAACGCC 4295  
 QY 882 ----- 882  
 Db 4296 CTCACCAAGACCTTGGCTTCGGGACCCGCTGTATGGAAGGAACCCCAATCCAGACAG 4355  
 QY 883 ----- 883  
 Db 4356 CCTGTCAGGAGGAGGAGAGTGGACCACTGCCCCCAGTTCCC---CAGACCATCATG 4412  
 QY 898 AspLeuTyr----- 900  
 Db 4413 GACCTCTTCAGNATGGGAATGGACATGCGAGNACCCCTTCACCTGCATGCCAGTGTAGC 4472  
 QY 901 -----PheLeuLysProGly-----Asp 906  
 Db 4473 AGCGACAAAATCAAGAAGATGCTGCTGTGTGTCCTCCCGAGGGGAGGGGGCTGCTCCT 4532  
 QY 907 LysProHisLysTyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAspIleSer 926  
 Db 4533 CCACAAAGAAAACAAAACACTGCGATATCCTTCAGGACTGACAGGAAGAACTTCG 4592  
 QY 927 Asp-----LeuIleSerPhePheThrSerGlnAsnIleMet 938  
 Db 4593 GATTATCTGCGTGAAGACGTATGTGCAGATCATAGCCAAAAGCTTAAAGAACAGATCTGG 4652  
 QY 939 ValThr----- 940  
 Db 4653 GTGAATGAGTTTGGTANGCGGCTTTTCCTGGGTGTCAGTAACTCAAGCACTTCT 4712  
 QY 941 -----MetIleAsnAspSer-----AspTyrValSerValAlaPro 952  
 Db 4713 CCGAGTCAAGAAGTTAATGATGCCATCAACAAATGAAGAAACACCTAAAGTGGCCAG 4772  
 QY 953 HisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyrVal----- 967  
 Db 4773 GACAGTTCTCGACATGATTTCTCAACAGCTTGGGAAGATTATGACAGACTGCGACACC 4832  
 QY 968 -----PheAlaAlaValPheAsnSerThrMetValTyrSerLeuProIleLeuVal 984  
 Db 4833 AGAATAATGTCAAGGTGTGTTCAATACACAGGCTGGCATGCAATCAGTCTTTCTCGT 4892  
 QY 985 AsnIleIleSerAsnTyrTyrLeu----- 992  
 Db 4893 AATGTCATCAACATGCTTCTCCGGGCCAACCTGCAAAAGGGAGAGAACCTAGCCAT 4952  
 QY 993 -----TyrHisLeuAsnValThrGluThrIleGlnIleTyrPser 1005  
 Db 4953 TATGGAAATTACTGTTTCATCATCCCTGAACTCACCAGCAG---CAGCTCTCAGAG 5009  
 QY 1006 ThrProPhePheGlnIleValThrAspIleValPheLysIleGluLeuTyrPheGlnAla 1025  
 Db 5010 GTGGCTCTGATGACCATCATGATGATGCTGCTGTGTCATCTGTGTCTCTTT----- 5063  
 QY 1026 AlaLeuLeuGlyIleIleValThrAlaMetProPyrPheAlaMetGluAsnAlaGlu 1045  
 Db 5064 -----GCAATGTCTCTGCTCCAGCCAGCTTTGCTGATTCCTGATCCAG 5108  
 QY 1046 AsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAlaTyr 1065  
 Db 5109 GAGCGGGTCAGCAAGCAAAACACCTGCAGTTTCATCATGTCAGTGAAGCCTGTCATCTAC 5168  
 QY 1066 TrpIleGlyGlnAlaValAspIleProLeuPhePheIleIleLeuIleLeuMetLeu 1085  
 Db 5169 TGGCTCTCAATTTTGTCTGGATATGTCATTCATGTTGCTCCCTGCCACACTGGTCAT 5228  
 QY 1086 GlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAlaVal 1105

Db 5229 ATCATCTTCATCTGCTTCCAGCAGAAGTCTCTATGTGCTCCACCAATCTGCTGTGCTA 5288  
 QY 1106 ValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSerPhe 1125  
 Db 5289 GCCCTTCTACTTTTGTCTGATGGTGGTCAATCACACACTCTCTCATGTACCCAGCCTCTTT 5348  
 QY 1126 ThrPheLysLysIleLeuAsnThrLysGluPheTyrPheIleTyrSerValAlaAla 1145  
 Db 5349 GTGTTTCAAG-----ATCCCCAGC 5366  
 QY 1146 LeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyr-----ThrIleAla 1163  
 Db 5367 ACAGCCTATGTGCTCACCAGCGTGAACCTCTTCATTGGCATTAAATGCGACGCTGGCC 5426  
 QY 1164 Thr-----IleLeuHis 1167  
 Db 5427 ACCTTTGTGTGAGCTGTTCCCGCAATAAGCTGAATAATATCAATGATATCTCTGAAG 5486  
 QY 1168 TyrAlaPheCysIleIleIleProIleTyrProLeuLeuGlyCysLeuIleSerPheIle 1187  
 Db 5487 TCCGTGTTCC---TTGATCTTCCACATTTTGGCTGGGACGAGGCTCATCGACATGGTG 5543  
 QY 1188 Lys----- 1188  
 Db 5544 AAAAACCCAGGCAATGGCTGATGCCGTGAAAGGTTGGGAGAAATCGCTTTGTGTGCCA 5603  
 QY 1189 IleSerTrpLysAsnValArgLysAsnValAspThrTyrAsnProTrpAspArgLeuSer 1208  
 Db 5604 TTAICTTTGGGACTTGGTGGGACGAAACCTCTTCCGCC-----ATGGCC 5645  
 QY 1209 ValAlaValIleSerProTyrLeuGlnCysValLeuTrpIlePheLeuLeuGlnTyrTyr 1228  
 Db 5646 GTGGAAGGGTGGTGTCTCTCTCACTACTGTT-----CTGATCCAGTAC--- 5690  
 QY 1229 GluLysLysTyrGlyArgSerIleArgLysAspProPhePheArgAsnLeuSerThr 1248  
 Db 5691 -----AGATCTTC-----ATCAGGCC 5708  
 QY 1249 LysSerLysAsnArgLysLeuProGluProAspAsnGluAspGluAspVal 1268  
 Db 5709 AGACTGTAAATGCAAGCTA-----TCTCTCTGAAT---GATGAAGATGAAGATGTG 5759  
 QY 1269 LysAlaGluArgLysValLysValLysGluLeuMetGlyCysGlnCysCysGluGluLysPro 1288  
 Db 5760 AGCGGGGAAACACAGAGAATTTCTTGATGGTGGAGGC-----CAGAATGAC 5804  
 QY 1289 SerIleMetValSerAsnLeuHisLysGluTyrAspAspLysLysAspPheLeuLeuSer 1308  
 Db 5805 ATCTTAGAAATCAAGGAGTTGACGAGATATAT----- 5837  
 QY 1309 ArgLysValLysLysValAlaThrLysTyrIleSerPheCysValLysLysGlyGluIle 1328  
 Db 5838 AGAAGGAAGCGAAGCCTGCTGTTGACAGGATTTGCTGGGCTCTCTCTGTTGATGTC 5897  
 QY 1329 LeuGlyLeuLeuGlyProAsnGlyValAlaGlyLysSerThrIleLeuAsnIleLeuValGly 1348  
 Db 5898 TTTGGCTCTCTGGGAGTTAATGGGCTGGAATAATCATCACTTTCAAGATGTTAAACAGGA 5957  
 QY 1349 AspIleGluProThrSerGlyGlnValPheLeuGlyAspTyrSerSerGluThrSerGlu 1368  
 Db 5958 GATACCACTGTTTACAGAGGAGATGCTTCTCT---AACAAAAATAGTATCTTATCAAC 6014  
 QY 1369 AspAspAspSerLeuLysCysMetGlyTyrCysProGlnIleAsnProLeuTrpProAsp 1388  
 Db 6015 ATCCATGAAGTATCATCAACATGGCTACTGCTCTGATTTGATGCCATCAGAGCTG 6074  
 QY 1389 ThrThrLeuGlnGluHisPheGluIleTyrGlyAlaValLysGlyMetSerAlaSerAsp 1408  
 Db 6075 TTTGACTGGGAGAAACACGTGGAGTCTCTTGGCTCTTTGAGAGGAGTCCCAGAGAAAGAA 6134  
 QY 1409 MetLysGluValIleSerArgIleThrHisAlaLeuAspLeuLysGluHisLeuGlnLys 1428

Db 2055 TATGAGAGGAGCGCTGAAGAGACCATGCGGATCATGGCGCTGGACACAGCATC 2114  
Qy 262 PheTrpLeuSerTrpValLeuLeuTyrThrSerLeuLeuPheLeuMet-----Ser 278  
Db 2115 CTTGCTTTAGCTGGTTCATT-----AGTAGCCTCATTCCTCTTGTGTGAGCGCTGGC 2168  
Qy 279 LeuLeuMetAlaValIleAlaThrAlaSerLeuLeuPheProGlnSerSerSerIleVal 298  
Db 2169 CTGCTAGTGTCTCATCTGAAGTAGGAACCTGCTG---CCCTACAGTATCCAGCGTG 2225  
Qy 299 IlePheLeuLeuPhePheLeuTyrGlyLeuSerSerValPhePheAlaLeuMetLeuThr 318  
Db 2226 GTGTTTGTCTTCCTGCTGCGGTGTGTGCTGGTGAACAATCTCTGAGTATGAC 2285  
Qy 319 ProLeuPheLeuLeuSerIleVal-----GlyIleValGluPhePheVal 334  
Db 2286 ACACCTCTTCCAGAGCCACCTGGCAGCAGCTGTGGGGGCATCATCTACTTACGCTG 2345  
Qy 335 ThrValAlaPhe-----GlyPheIleGlyLeuMetIleLeu 347  
Db 2346 TACCTGCCCTACCTCTGTGTGTGGCATGGCAGGACTACGTGGGCTTCACACTCAAGATC 2405  
Qy 348 IleGluSerPheProLysSerLeuValTrpLeuPheSerProPheCysHisCysThrPhe 367  
Db 2406 TTCGCTAGC-----CTGCTGTCTCCT-----GTGGCTTTT 2435  
Qy 368 ValIleGlyIleAlaGlnValMetHisLeuGluAspPheAsnGluGlyAlaSerPheSer 387  
Db 2436 GGGTTTGGCTGTGAGTACTTTGGCTTTTGGAGGAGGAGGATGGAGTGGAGTGGGAC 2495  
Qy 388 AsnLeuThrAlaGlyPro-----TyrProLeuIleIleThrIleLeuMet 402  
Db 2496 AACCTTTTGAGAGTCTCTGTGGAGGAGATGGCTTCAATCTCACCATTCGGTCTCCATG 2555  
Qy 403 LeuThrLeuAsnSerIlePheTrpValLeuAlaValTyrLeuAspGlnValIlePro 422  
Db 2556 ATGCTGTTTGACACCTTCTCTATGGGGTATGACCTGCTGATACATGAGCGTCTCTTCCA 2615  
Qy 423 GlyGluPheGlyLeuArgSerSerLeuTyrPheLeuLeuProSerTrp----- 440  
Db 2616 GGCAGTAGCGAATTCCTGAGCGCTGCTGTTATTTCTTGACCAAGTCTTACTGTTGGC 2675  
Qy 441 -----SerLysSerLysArgAsnTyrGluGluLeuSerGluGlyAsn 454  
Db 2676 GAGAAAGTGTAGAGAGAGCCACCTGCTTCCACAGCAGAGATATCAGAA----- 2729  
Qy 455 ValAsnGlyAsnIleSerPheSerGluIleLeuGluProValSerSerGluPheValGly 474  
Db 2730 -----ATCTGCTATGAGGAG-----GAACCCACCCACTTG----- 2759  
Qy 475 LysGluAlaIleArgIleSerGlyIleGlnLysThrTyrArgLysLysGlyGluAsnVal 494  
Db 2760 AGCTGGGGGTGCTCAATTCAGAACCTGGTAAAGTCTACCGA---GATGGGATGAGGTG 2816  
Qy 495 GluAlaLeuArgAsnLeuSerPhePheIleTyrGluGlyGlnIleThrAlaLeuLeuGly 514  
Db 2817 ---GCTGTGATGGCGCTGCACTGAATTTTATGAGGCGCAGATCACCTCTCTCTGGGC 2873  
Qy 515 HisSerGlyThrGlyLysSerThrLeuMetAsnIleLeuCysGlyLeuCysProSer 534  
Db 2874 CACATGGAGCGGGAGACGACCACTGTCATCTGACCGGTGTTCCTCCCGGACC 2933  
Qy 535 AspGlyPheAlaSerIleTyrGlyHisArgValSerGluIleAspGluMetPheGluAla 554  
Db 2934 TCGGGCACCGCTTACATCTGGGAAAAGACATTGCG-----TCTGAGATGAGCACCATC 2987  
Qy 555 ArgLysMetIleGlyIleCysProGlnLeuAspIleHisPheAspValLeuThrValGlu 574  
Db 2988 CGCAGAACTGGGGGTCTGCTCCCGCAGCATACAGTCTGTTTGAATGCTGACTCTGAA 3047  
Qy 575 GluAsnLeuSerIleLeuAlaSerIleLysGlyIleProAlaAsnAsnIleIleGlnGlu 594  
Db 3048 GAACACATCTGTTCTATGCCCGCTTGAAGGGCTCTCTGAGAGCAGCTGAGGCGGAG 3107

Qy 595 ValGlnLysValLeuLeuAspLeuAspMetGlnThrIleLys---AspAsnGlnAlaLys 613  
Db 3108 ATGAGCAGATGGCCCTGGATGTGGTTTGGCCATCAGCAGCTGAAACCAACACAGC 3167  
Qy 614 LysLeuSerGlyGlyGlnLysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsnPro 633  
Db 3168 CAGCTGTGAGTGAATGACAGAGAAAGCTATCTGTGGCCTTGGCTTGTGCGGGATCT 3227  
Qy 634 LysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIleVal 653  
Db 3228 AAGTTGTCTATCTGGATGAACCCACAGCTGGTGTGACCTTACTCCCCAGGGGAATA 3287  
Qy 654 TrpAsn---LeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPhe 672  
Db 3288 TGGGAGCTGCTGTGAATATACCGCAA---GGCGGCACCATTTATCTCTACACACCAC 3344  
Qy 673 MetAspGluAlaPheIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLys 692  
Db 3345 ATGATGAAGCGGACGCTCTGGGGGACAGGATGCCATCATCTCCCATGGGAAGCTGTC 3404  
Qy 693 CysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyTyrArgLeuSerMet 712  
Db 3405 TGTGTGGCTCTCTCCCTGTTTCTGAAGAACAGCTGGGAACAGGCTACTACTGACCTTG 3464  
Qy 712 ----- 712  
Db 3465 GTCAAGAAAGATGTGGAAATCCTCCTCAGTTCTCTGCAGAAACAGTAGTAGCTGTGTC 3524  
Qy 713 TyrIleAspLysTyrCysAla----- 719  
Db 3525 TACCTGAAAAGGAGGACAGTGTTCACAGAGCAGTTCTGATGCTGGCGCTGGGACGGAC 3584  
Qy 720 -----ThrGluSerLeuSerSerLeuValLysGlnHis 730  
Db 3585 CATGAGAGTGACACGCTGACCATCGATGCTCTGCTATCTCCACCTCATCAGAGCAT 3644  
Qy 731 IleProGlyAlaThrLeuLeuGlnAsnAspGlnGlnLeuValTyrSerLeuProPhe 750  
Db 3645 GTGTCTGAACCGCGCTGGTGGAGACATAGGGCATGAGCTGACCTATGTGCTGCCATAT 3704  
Qy 751 LysAspMetAspLys-----PheSerGlyLeuPheSerAlaLeuAspSerHis---Ser 767  
Db 3705 GAGCTCTGAAGGAGGAGCGCTTTGTGAACTCTTTCATGAGATGATGATCGCGCTCTCA 3764  
Qy 768 AsnLeuGlyValIleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLys 787  
Db 3765 GACCTGGGCATTTCTAGTTATGGCATCTCAGAGACGACCTGGAGAAATATTCTCTCAAG 3824  
Qy 788 LeuGluValGluAlaGluIleAsp----- 795  
Db 3825 GTGCCCGAAGAGAGTGGGGTGGATGCTGAGACCTCAGATGGTACCTGCCAGCAGACGA 3884  
Qy 796 -----GlnAlaAspTyrSerValPheThrGln----- 804  
Db 3885 AACAGGCGGCGCTTCGGGGCAACAGCAGAGCTGTCTTCGCCCGTTCACTGAGATGATGCT 3944  
Qy 805 ---GlnProLeuGluGluMetAspSerLysSerPheAspGluMetGluGlnSerLeu 823  
Db 3945 GCTGATCCAAATGATCTGACATAGACCCAGATCC-----AGAGAGACAGACTTG 3995  
Qy 824 LeuIleLeuSerGluThrLysAlaSerLeu---ValSerThrMetSerLeuTrpLysGln 842  
Db 3996 CTCAGTGGGATGGATGGCAAGGGTCTCTACAGGTGAAGGCTGGAAACTTACACAGCAA 4055  
Qy 843 GlnMetTyrThrIleAlaLysPheHisPhePheThrLeuLysArgGluSerLysSerVal 862  
Db 4056 CAGTTTGTGGCCCTTTTGTGGAGAGACTGCTAATTGCGCAGAGGAGTGGGAAGGATTT 4115  
Qy 863 ArgSerValLeuLeuLeuLeuIlePhePheThrValGlnIlePheMetPheLeuVal 882  
Db 4116 TTTGCTCAGATTGCTTGGCCAGCTGTGTTGTGCTGATTCGCCCTTGTGTTTCACCTGATC 4175



GENERAL INFORMATION:	
; APPLICANT: Hayden, Michael R.	
; APPLICANT: Brooks-Wilson, Angela R.	
; APPLICANT: Pimstone, Simon N	
; TITLE OF INVENTION: METHODS AND REAGENTS FOR MODULATING	
; TITLE OF INVENTION: CHOLESTEROL LEVELS	
; FILE REFERENCE: 50110/002005	
; CURRENT APPLICATION NUMBER: US/09/526,193A	
; CURRENT FILING DATE: 2000-03-15	
; PRIOR APPLICATION NUMBER: 60/124,702	
; PRIOR FILING DATE: 1999-03-15	
; PRIOR APPLICATION NUMBER: 60/138,048	
; PRIOR FILING DATE: 1999-06-08	
; PRIOR APPLICATION NUMBER: 60/139,600	
; PRIOR FILING DATE: 1999-06-17	
; PRIOR APPLICATION NUMBER: 60/151,977	
; PRIOR FILING DATE: 1999-09-01	
; NUMBER OF SEQ ID NOS: 287	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 2	
; LENGTH: 7860	
; TYPE: DNA	
; ORGANISM: Homo sapiens	
US-09-526-193A-2	
Alignment Scores:	
Pred. No.: 4,41e-145 Length: 7860	
Score: 1479.00 Matches: 490	
Percent Similarity: 42.75% Conservative: 318	
Best Local Similarity: 25.93% Mismatches: 624	
Query Match: 17.55% Indels: 458	
DB: 4 Gaps: 65	
US-10-090-458-5 (1-1642) x US-09-526-193A-2 (1-7860)	
QY	89 SerSerIleMetGlnLysValSerThrAspHisLeuProaspVallelleThrGluGlu 108
DB	1569 TCTCGCTTCATGGAGTGTCAACCTGACAGCTA---GAACCCATAGACAGAGTC 1625
QY	109 TyrThrAsnGluLysGluMetLeuThrSerSerLeuSerLysProSerAsnPheValGly 128
DB	1626 TGGCTCATCAACAGTCCATG-----GAGCTCTGATGAGAGGAGTTCGGGTGGT 1679
QY	129 ValValPhe-----LysAspSerMetSerTyrGlu 138
DB	1680 ATTGTGTTCACTGGATTACTCCAGGAGCATTGAGTCCCATCATGTCAAGTACAAG 1739
QY	139 LeuArgPhePheProaspMetIle-----ProValSerSerIleTyrMetAsp 154
DB	1740 ATCCGAATGGACATGTGCAATGTGGAGAGGACAAATAAATCAAGATGGGTACTGGAC 1799
QY	155 -----SerArgAlaGlyCySerLysSerCysGluAlaAlaGlnTyrTrpSerSergly 172
DB	1800 CCTGGTCTCTGAGCT-----GACCCCTTTGAGGACATGCGGTACGTCGGGGGGC 1850
QY	173 PheThrValLeuGlnAlaSerIleAspAlaAlaIleIleGlnLeuLysThrAsnValSer 192
DB	1851 TTCGCCTACTTGCAGGATGTGGTGGACAGCAATCATCAGGTGTGACGGGACAC--- 1907
QY	193 LeuTrpLysGluLeuGluSerThrLysAlaValIleMetGlyGluThrAlaValValGlu 212
DB	1908 -----GAGAGAAACTGGTGTCTATATGCAACAGATG----- 1940
QY	213 IleAspThrPheProArgGlyValIleLeuIleTyrLeuValIleAlaPheSerProPhe 232
DB	1941 -----CCCTATCCCTGTACGTTGATGACATCTTTCTGGGGGTGATGAGCCGGTCAATG 1994
QY	233 GlyTyrPheLeuAlaIle-----HisIleVal 241
DB	1995 CCCCTCTTCATGACGCTGGCTGGATTACTCACTGGCTGTGATCATCAAGGGCAUCGTG 2054
QY	242 AlaGluLysGluLysLysIleLysGluPheLeuLysIleMetGlyLeuHisAspThrAla 261

Db 2048 ATCTGGAGTCTTCTTTCAGCGGCAAGAAAGTGCACCGACCATCGTGTGACCAACCTTC 2107  
Qy .....  
Db 673 MetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLys 692  
Qy .....  
Db 2108 ATGGACGAGGCTGACCTGTCTGGGAGACCGCATCGCCATCATGGCCAAAGGGAGGTGCAG 2167  
Qy .....  
Db 693 CysValGlySerSerMetPheLeuLysSerLysTrpGlyIleGlyLysArgLeuSerMet 712  
Qy .....  
Db 2168 TGTCTGGGTCTCTGCTGCTCTCAAGCAGAAATACGGTCCGCTATCATAGAGCTG 2227  
Qy .....  
Db 713 TyrIleAspLysTyrCysAlaThrGluSerLysSerLysValGlnHisLeuPro 732  
Qy .....  
Db 2228 GTCAAGAGCGGCACCTGCACACCGGAAGACATCTCCAGCTGTCTCCACCAACGCTGCC 2287  
Qy .....  
Db 733 GlyAlaThrLeuLeuGlnGlnAsnAspGlnLeuValTyrSerLeuProPheLysAsp 752  
Qy .....  
Db 2288 AACGCCACGCTGAGAGACGCTGGGGCGAGCTGTCTTTCATCTCCAGAGAGAC 2347  
Qy .....  
Db 753 MetAspLysPheSerGlyLeuPheSerAlaLeuAsp---SerHisSerAsnLeuGlyVal 771  
Qy .....  
Db 2348 ACGCACAGGTTTGAAGCTCTCTTTGTAAACTGGAGAAGAGCAGAAAGAGCTGGGCATT 2407  
Qy .....  
Db 772 IleSerTyrGlyValSerMetThrLeuGluAspValPheLeuLys--- 787  
Qy .....  
Db 2408 GCCAGCTTTGGGCACTCCATCACCATGGAGGAAGTCTTCTTCTGGGTTCGGGAAGCTG 2467  
Qy .....  
Db 788 ----- 793  
Qy .....  
Db 2468 GTGACAGCAGTATGACATCCAGGCCATCCAGCTCCCTGCTGCTGACGACGACGAG 2527  
Qy .....  
Db 794 IleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluMetAspSer 813  
Qy .....  
Db 2528 AGCGCGGCCAGCAGCTGGGCTGTG-----GACAGC 2557  
Qy .....  
Db 814 LysSerPheAspGluMetGluGlnSer-----LeuLeuIleLeuSerGlu 828  
Qy .....  
Db 2558 AACCTGTGGGGCCATGGACCCCTCCGAGCGCATTTGGAGCCCTCATCGAGGAGGAGCGC 2617  
Qy .....  
Db 829 ThrLysAlaSerLeuValSerThrMetSerLeuTrpLysGlnMetTyrThrIleAla 848  
Qy .....  
Db 2618 ACGGCTCAAGCTCAACACTGGGCTCGCCCTGCACTGCAGCAATCTGGGCCCAT--- 2674  
Qy .....  
Db 849 LysPheHisPheThrLeuLys-----ArgGluSerLysSerValArg 863  
Qy .....  
Db 2675 -----TTCTGAGAAAGCGCGCATACAGCTGGCGGAGTGAATATGTGGCG 2722  
Qy .....  
Db 864 Ser-----ValLeuLeuLeuLeuLeuIlePhePheThr 874  
Qy .....  
Db 2723 GCACAGGCTCGTGTGCTCTGACCTGCGTCACTGCGCTGCGCTGCGCTCACTCACTCC 2782  
Qy .....  
Db 875 ValGlnIlePhe-----MetPheLeuValHisSerPheLysAsnAlaVal 890  
Qy .....  
Db 2783 TCGAGCTCTTCAGAGACCCCATGCTGAGGCTGACCTTGGCGAGTACGCGCAACCGTC 2842  
Qy .....  
Db 891 ValProIleLysLeuValProAspLeuTyrPheLeuLysPheProGlyAspLysProHisLys 910  
Qy .....  
Db 2843 GTGCCC-----TTCTCAGTTCCCGGGACCTCCACGCTGGGT 2878  
Qy .....  
Db 911 TyrLysThrSerLeuLeuLeuGlnAsnSerAlaAspSerAsp----- 924  
Qy .....  
Db 2879 CACAGCTGTACAGACATCTGAAGACGCACTGCAGGCTGAGGAGCAGGAGCCCGCGAG 2938  
Qy .....  
Db 925 ---IleSerAspLeuIleSerPhe-----Phe 932  
Qy .....  
Db 2939 GTGCTCGGTGACCTGGAGGATCTTGATCTTCAGGGCTTCTGTGGAGGGGGCGGCTTT 2998  
Qy .....  
Db 933 ThrSerGlnAsnIleMetValThrMetIleAsnAspSerAspTyrValSerValAlaPro 952  
Qy .....  
Db 2999 AATGAGCGGCTGTGGCAGCGCTCCTTCAGAGAT-----GTGGGA 3040  
Qy .....  
Db 953 HisSerAlaAlaLeuAsnValMetHisSerGluLysAspTyr-----Val 967  
Qy .....  
Db 967 -----Val 967

Db 3041 GAGCGCAGGTCGTCAACGCGCTTGTTCACAAACAGCGGTACCACTCTCCAGCCACTGCC 3100  
Qy .....  
Db 968 PheAlaAlaValPheAsnSerThrMetValTyrSerLeu-----ProIleLeuVal 984  
Qy .....  
Db 3101 CTGCGCGTCGCGGACAAC-----CTTCTGTTCAGCTGTCTGCGGGCTCTACGCGCTCC 3154  
Qy .....  
Db 985 AsnIleIleSerAsnTyrTyrLeuTyrHisLeuAsnValThrGluThrIleGlnIleTrp 1004  
Qy .....  
Db 3155 ATTGTGCTCTCAACTTC-----CCCCAGCCCGGAGCGCCTGCAGCTGCC 3202  
Qy .....  
Db 1005 SerThrProPhePheGlnGluIleThrAspIleValPheLysIleGluLeuTyrPheGln 1024  
Qy .....  
Db 3203 AAGACACGAGTTTAAAGAGGCGCGGAGGATTCGACATTCCTCCCAACCTGCTCTTCGCC 3262  
Qy .....  
Db 1025 AlaAlaLeuLeuGlyIleIleValThrAlaMetProTyrPheAlaMetGluAsnAla 1044  
Qy .....  
Db 3263 ATGGCATCTTCGGC-----AGCAGCTTCTCCATCTCCGCGCTC 3301  
Qy .....  
Db 1045 GluAsnHisLysIleLysAlaTyrThrGlnLeuLysLeuSerGlyLeuLeuProSerAla 1064  
Qy .....  
Db 3302 AGCGAGAGGCGCGTGCAGGCCAAGCATGTGCAGTTTGTGAGTGGAGTCCACGTCGCCAGT 3361  
Qy .....  
Db 1065 TyrTrpIleGlyGlnAlaValAspIleProLeuPhePheIleIleLeuLeuMet 1084  
Qy .....  
Db 3362 TTCCTGGCTCTGCTCTGCTGTGGACCTCATCTCTCTCTCATCCAGCTCTGCTGTG 3421  
Qy .....  
Db 1085 LeuGlySerLeuLeuAlaPheHisTyrGlyLeuTyrPheTyrThrValLysPheLeuAla 1104  
Qy .....  
Db 3422 CTGCTGTGTGTTTAAAGGCTTCGACGTGCTGCTTCACGCGGGAGCGCCACATGGCTGAC 3481  
Qy .....  
Db 1105 ValValPheCysLeuIleGlyTyrValProSerValIleLeuPheThrTyrIleAlaSer 1124  
Qy .....  
Db 3482 ACCCTGCTGTCTCTGCTCTACGGTGGCCATCATCCCTCATGTACTCATGATGAC 3541  
Qy .....  
Db 1125 PheThrPhe-----LysLysIleLeuAsnThrLysGluPheTrpSerPheIleTyr 1141  
Qy .....  
Db 3542 TTCTTCTTCTTGGGGCGGCCACTGCTTACAGAGGCTGACCATCTTCAACATCTGTCA 3601  
Qy .....  
Db 1142 SerValAlaAlaLeuAlaCysIleAlaIleThrGluIleThrPhePheMetGlyTyrThr 1161  
Qy .....  
Db 3602 GGCATCGCCACTTCTGTGTGTGTCACCATCATGCGCATCCAGCTGTAAACTGGAAGAA 3661  
Qy .....  
Db 1162 IleAlaThrIleLeuHisTyrAlaPheCysIleIleIleIleProIleTyrProLeuLeuGly 1181  
Qy .....  
Db 3662 CTTTCCAAAACCTTGGATCAGTGTTC---CTGCTGTGCTGCCAACCCAC---TGCTGGGG 3715  
Qy .....  
Db 1182 CysLeuIleSerPheIleLysIleSerTrpLysAsnValArg----- 1195  
Qy .....  
Db 3716 ATGCGAGTCAGCAGTTCCTACGAGAACTACGAGACGCGGAGGTACTGCACCTCTCCGAG 3775  
Qy .....  
Db 1196 -----LysAsnValAspThrTyrAsnProTrp 1204  
Qy .....  
Db 3776 GTCGCGCCCACTACTGCAAGAAATATAACATCCAGTACGAGGAGAACTTCTATGCTCGG 3835  
Qy .....  
Db 1205 AspArgLeuSerValAla-----ValIleSerProTyrLeuGlnCysValLeuTrpIle 1222  
Qy .....  
Db 3836 AGCGCCCGGGGTGCGCGGCTTGTGCGCTTCCATGCGCCCTCAGGCTGCGCTACCTC 3895  
Qy .....  
Db 1223 PheLeuGlnIleTyrTyrGlyLysTyrGlyArg-----Ser 1236  
Qy .....  
Db 3896 ATCTGCTCTTCTTCATCATGAGACCAACCTGCTCAGAGACTCAGGGGCACTCTCTGGGCC 3955  
Qy .....  
Db 1237 IleArgLysAspProPhePheArgAsnLeuSerThrLysSerLysAsnArgLysLeuPro 1256  
Qy .....  
Db 3956 CTCGGAGGAGCGGACACTGACAGAAATATACACC-----CGGATGCTC 4000  
Qy .....  
Db 1257 GluProProAspAsnGluAspGluAspValLysAlaGluArgLeuLysValLys 1276  
Qy .....  
Db 4001 GTGCTTCT-----GAGACCAAGATAGCGGACGAGACCCGATCTCTG 4048  
Qy .....  
Db 1277 GluLeuMetGlyCysGlnCysCysGluLysProSer----- 1289  
Qy .....  
Db 4049 -----GCCCGGAGCGGACTCCTCTCTCCACACA 4078





593	QY	GlnGluValGlnIbysValLeuLeuAspLeuAspMetGlnThrThrIleIysAspAenGlnAla	612
1868	DB	GAAGAAGTCAAGCAGATGCTGCACATCGGCTGGAGACAAGTGAAGTCACTCAGGAGC	1927
613	QY	LysLysLeuSerGlyGlyGlnIbysArgLysLeuSerLeuGlyIleAlaValLeuGlyAsn	632
1928	DB	CGCTTCTTGAGCGGGGCATGAGCGCAAGCTCTCCATCGGCATCGCCTCATCGCAGGC	1987
633	QY	ProLysIleLeuLeuLeuAspGluProThrAlaGlyMetAspProCysSerArgHisIle	652
1988	DB	TCCAAAGTGTGATCTGAGCAGACCCACCTCGGCGATGACGCCATCTCCAGAGAGGCC	2047
653	QY	ValTrpAsnLeuLeuLysTyrArgLysAlaAsnArgValThrValPheSerThrHisPhe	672
2048	DB	ATCTGGGATCTTCTTCAGCGCGCAGAAAGTGACCGCACCATCTGCTGTGACCAACCCACTTC	2107
673	QY	MetAspGluAlaAspIleLeuAlaAspArgLysAlaValIleSerGlnGlyMetLeuLys	692
2108	DB	ATGGACGAGGCTGACCTGCTGGGAGACCGCATCGCCATCATGGCCAAAGGGGAGCTGCAG	2167
693	QY	CysValGlySerSerMetPheLeuLysSerLysTrpIlyIleGlyTyrArgLeuSerMet	712
2168	DB	TGTCGGGTCTCGCTGTTCTTCAAGCAGAAATACGGTCCGGCTATCATCAACCGCTG	2227
713	QY	TyrIleAspLysTyrCysAlaThrGluSerLeuSerSerLeuValIysGlnHisIlePro	732
2228	DB	GTGAAGAGCGGCACCTCAACCCGGGAAGCATCTCCAGCTGTGTCCACCAACGAGTCCC	2287
733	QY	GlyAlaThrLeuLeuGlnGlnAsnAspGlnLeuValTyrSerLeuProPheLysAsp	752
2288	DB	AACGCCACGCTGGAGAGCAGCGTGGGGCCGAGCTGTCTTTCATCTCTCCAGAGAGC	2347
753	QY	MetAspLysPheSerGlyLeuPheSerAlaLeuAsp--SerHisSerAsnLeuGlyVal	771
2348	DB	ACGCACAGGTTTGAAGTCTCTTTGCTAACTGGAGAAGACGAGAAGAGTGGGCATT	2407
772	QY	IleSerTyrGlyValSerMetThrThrLeuGluAspValPheLeuLys-----	787
2408	DB	GCCAGCTTTGGGGCATCCATCACCATGGAGAGAGTCTCTCTCGGTGGGGAAGCTG	2467
788	QY	-----LeuGluValGluAlaGlu	793
2468	DB	GTGACACAGCATATGGACATCCAGGCCATCCAGCTCCCTGCCCTGCAGTACCAGCAG	2527
794	QY	IleAspGlnAlaAspTyrSerValPheThrGlnGlnProLeuGluGluMetAspSer	813
2528	DB	AGCGCGCCAGCGACTGGGTGTG-----GACAGC	2557
814	QY	LysSerPheAspGluMetGluGlnSer-----LeuLeuLeuLeuSerGlu	828
2558	DB	AACCTGTGGGGCCATGGACCCCTCCAGCGCATTCGAGCCCTCATCGAGGAGGAGCGC	2617
829	QY	ThrIysAlaSerLeuValSerThrMetSerLeuTrpLysGlnGlnMetTyrThrIleAla	848
2618	DB	ACCCTGTCAAGCTCAACACTGGGCTGGCCCTGCACCTGCCACGAATCTCTGGGCCATG	2674
849	QY	LysPheHisPheThrLeuLys-----ArgGluSerLysSerValArg	863
2675	DB	-----TTCCTGAAGAGCGCGCATACAGCTGCCTGCGAGTGGAAAATGTTGGCG	2722
864	QY	Ser-----ValLeuLeuLeuLeuLeuLeuPhePheThr	874
2723	DB	GCACAGTCTCTGGTCTGTGACCTGCGTCACTCCCTGGCCCTCTCTGGCCATCACTACTCC	2782
875	QY	ValGlnIlePhe-----MetPheLeuValHisHisSerPheLysAsnAlaVal	890
2783	DB	TCGAGCTCTTCGACGACCCCATCTGAGGCTGACCTTGGCGGAGTACGGCAGAACCGCTC	2842
891	QY	ValProIleLysLeuValProAspLeuTyrThrPheLeuLysProGlyLysPheHisLys	910
2843	DB	GTGGCC-----TTCACAGTTCCTCCGGCACTCCACGAGTGGT	2878
911	QY	TyrLysThrSerLeuLeuGlnAsnSerAlaAspSerAsp-----	924

